



การพัฒนาระบบประมวลผลภาพสำหรับการตรวจสอบเส้นดำในการผลิตกุ้งแช่แข็ง  
Development of an Image Processing System for Detecting of Vein in  
Frozen Shrimp Process

นวพร ธนะสาร  
Nawaporn Thanasarn

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต  
สาขาวิชาวิศวกรรมอุตสาหกรรมและระบบ  
มหาวิทยาลัยสงขลานครินทร์

A Thesis Submitted in Partial Fulfillment of the Requirements for the  
Degree of Master of Engineering in Industrial and Systems Engineering  
Prince of Songkla University

2561

ลิขสิทธิ์ของมหาวิทยาลัยสงขลานครินทร์



การพัฒนาระบบประมวลผลภาพสำหรับการตรวจสอบเส้นดำในการผลิตกุ้งแช่แข็ง  
Development of an Image Processing System for Detecting of Vein in  
Frozen Shrimp Process

นวพร ธนะสาร

Nawaporn Thanasarn

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมอุตสาหกรรมและระบบ

มหาวิทยาลัยสงขลานครินทร์

A Thesis Submitted in Partial Fulfillment of the Requirements for the  
Degree of Master of Engineering in Industrial and Systems Engineering

Prince of Songkla University

2561

ลิขสิทธิ์ของมหาวิทยาลัยสงขลานครินทร์



ขอรับรองว่า ผลงานวิจัยนี้มาจากการศึกษาวิจัยของนักศึกษาเอง และได้แสดงความขอบคุณบุคคลที่มีส่วนช่วยเหลือแล้ว

ลงชื่อ.....

(ผู้ช่วยศาสตราจารย์ ดร.สุภาพรรณ ไชยประพัทธ์)

อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

ลงชื่อ.....

(ดร.เกรียงไกร ไวยกาญจน์)

อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม

ลงชื่อ.....

(นางสาวนวพร ณะสาร)

นักศึกษา

ข้าพเจ้าขอรับรองว่า ผลงานวิจัยนี้ไม่เคยเป็นส่วนหนึ่งในการอนุมัติปริญญาในระดับใดมาก่อน และ  
ไม่ได้ถูกใช้ในการยื่นขออนุมัติปริญญาในขณะนี้

ลงชื่อ.....

(นางสาวนภาพร ณะสาร)

นักศึกษา

ชื่อวิทยานิพนธ์	การพัฒนาระบบประมวลผลภาพสำหรับการตรวจสอบเส้นดำในการผลิต กุ้งแช่แข็ง
ผู้เขียน	นางสาวนวพร ธนะสาร
สาขาวิชา	วิศวกรรมอุตสาหการและระบบ
ปีการศึกษา	2560

### บทคัดย่อ

เส้นดำเป็นระบบทางเดินอาหารของกุ้งที่เต็มไปด้วยก้อนกรวด ททราย และตะกอนซึ่งสามารถมองเห็นเป็นเส้นสีดำยาวไปตามลำตัวได้อย่างง่าย การปรากฏของเส้นดำเป็นข้อจำกัดในหลายผลิตภัณฑ์ เช่นเดียวกับกับในมาตรฐานผลิตภัณฑ์กุ้งสดและแช่แข็งของประเทศสหรัฐอเมริกาที่ระบุว่ากุ้งที่มีเส้นดำหลงเหลือเกินกว่ามาตรฐานกำหนดในผลิตภัณฑ์กุ้งที่ต้องถูกกำจัดเส้นดำถือว่าขาดคุณสมบัติ ปัจจุบันขั้นตอนการกำจัดและการตรวจสอบเส้นดำใช้คนเป็นแรงงานหลัก ก่อให้เกิดการขาดแคลนแรงงานผลักดันให้อุตสาหกรรมต้องพึ่งแรงงานต่างด้าว การตรวจสอบคุณภาพอัตโนมัติจึงเป็นเทคโนโลยีที่ไม่สามารถปฏิเสธได้อีกต่อไป งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อพัฒนาวิธีการตรวจจับกุ้งที่มีเส้นดำไม่ถูกต้องด้วยข้อมูลภาพ โดยใช้ภาพถ่ายกุ้งจำนวน 200 ภาพทดลองด้วยเครื่องมือประมวลผลภาพ ซึ่งประกอบด้วยการประมวลผลเบื้องต้นโดยการแปลงเป็นภาพระดับเทา การแยกพื้นที่ที่สนใจออกจากพื้นหลัง (ROI) และการสกัดคุณลักษณะที่สำคัญ โดยคุณลักษณะเหล่านี้ได้จากการการวัดรูปร่าง ได้แก่ พื้นที่ และความยาว และการวัดค่าพิคเซล ได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าต่ำสุด เปอร์เซ็นไทล์ 25, 50 และ 75 ในงานวิจัยนี้พบว่ากุ้งที่ไม่ผ่านเกณฑ์มาตรฐานถูกจำแนกด้วยเทคนิคการจำแนกที่แตกต่างกัน 2 วิธี คือ วิธี Discrimination Analysis (DA) และวิธีซัพพอร์ตเวกเตอร์แมชชีน (SVM) นอกจากนั้นมีการใช้เทคนิคการลดขนาดข้อมูล การวิเคราะห์องค์ประกอบหลัก (PCA) ในการเพิ่มประสิทธิภาพการจำแนก ซึ่งพบว่าเทคนิคดังกล่าวส่งผลต่อความถูกต้องในการทำนายเพียงเล็กน้อย โดยวิธี SVM เคอร์เนลฟังก์ชันเชิงเส้นเป็นวิธีที่ให้ความแม่นยำในการจำแนกสูงสุด ความสำเร็จของงานวิจัยนี้ไม่เพียงแต่เติมช่องว่างที่ยังหลงเหลือจากการศึกษาที่ผ่านมา แต่ยังเป็นที่ยืนยันถึงแนวโน้มของอนาคตที่ดีของการพัฒนาการตรวจสอบคุณภาพกุ้งแบบอัตโนมัติ

**Thesis Title**        Development of an Image Processing System for Detecting of Vein  
   in Frozen Shrimp Process

**Author**                Miss Nawaporn Thanasarn

**Major Program**    Industrial and Systems Engineering

**Academic Year**    2017

### ABSTRACT

The so-called “vein” in shrimps is a digestive tract filled with grit, sand and sediments, easily viewed as a dark string stretching along the abdomen. Presence of vein is restricted in various products, as detailed in the United States Standards for Grades of Fresh and Frozen Shrimp. Deveined shrimps with remnant vein longer than specified therein are considered disqualified. Currently, removing vein and discriminating improperly deveined shrimps are manually handled. Scarcity of local labor drives the industry to increasingly rely on uncertain supply of migrant workers. It has come to the point where necessity of automated quality inspection technologies cannot be denied anymore. This research aims to develop an image-based approach for detection of improperly deveined shrimps. Two hundred shrimp images were experimented by a sequence of image processing tools; preprocessing to gray scale images, segmenting the region of interest (ROI), and extracting significant features. These features include shape measurements, i.e., area and length, and pixel value measurements, i.e., average, standard deviation, minimum, 25th, 50th, and 75th percentiles. In this research, disqualified shrimps were identified by two different classification techniques: discrimination analysis (DA) and support vector machine (SVM). Merit of Principal Component Analysis (PCA), a dimensional reduction technique, on classification performance was investigated. However, it was found that in this research slight contribution of PCA was observed. Highest classification accuracy was obtained from the SVM with a linear kernel function. The success of this research not only fills a void left by past studies, but also assures a promising future of fully automated shrimp quality inspection development.

## กิตติกรรมประกาศ

ขอขอบพระคุณผู้ช่วยศาสตราจารย์ ดร.สุภาพรรณ ไชยประพัทธ์ ประธานกรรมการที่ปรึกษา ดร.เกรียงไกร ไวยกาญจน์ กรรมการที่ปรึกษาร่วม ที่กรุณาให้คำแนะนำแนวทาง และวิธีแก้ไขปัญหาในการทำวิทยานิพนธ์และการติดตามความก้าวหน้าของงานวิจัย รวมไปถึงการตรวจทานความถูกต้องของวิทยานิพนธ์ให้มีความถูกต้องสมบูรณ์

ขอขอบพระคุณรองศาสตราจารย์สมชาย ชูโณม รองศาสตราจารย์ ดร.พรชัย พลฤกษ์ภัทรานนท์ และผู้ช่วยศาสตราจารย์ ดร.ชุมพล ยวงใย กรรมการผู้ทรงคุณวุฒิ ที่กรุณาเป็นคณะกรรมการสอบป้องกันวิทยานิพนธ์

ขอขอบคุณคณะกรรมการสอบความก้าวหน้าวิทยานิพนธ์ทุกท่านที่ให้ข้อเสนอแนะและแนวทางในการแก้ไขปัญหา ซึ่งเป็นผลทำให้บรรลุวัตถุประสงค์ของวิทยานิพนธ์และเกิดความสมบูรณ์มากยิ่งขึ้น

ขอขอบคุณบุคลากรภาควิชาวิศวกรรมอุตสาหกรรม คณะวิศวกรรมศาสตร์ มหาวิทยาลัยสงขลานครินทร์ ทุกๆ ท่านที่ให้ความช่วยเหลือตั้งแต่เริ่มต้นจนกระทั่งการดำเนินการวิจัยสำเร็จลุล่วงไปด้วยดี และขอขอบคุณบัณฑิตคณะวิศวกรรมศาสตร์และบัณฑิตวิทยาลัย มหาวิทยาลัยสงขลานครินทร์ที่ให้ความสะดวกและคำแนะนำในเรื่องต่างๆ

ขอขอบคุณหน่วยบริหารงานวิจัย คณะวิศวกรรมศาสตร์ และบัณฑิตวิทยาลัย มหาวิทยาลัยสงขลานครินทร์ ที่สนับสนุนเงินทุนงานวิจัย และขอขอบคุณทีมงานของหน่วยบริหารงานวิจัย ที่ให้ความช่วยเหลือและคำปรึกษาในเรื่องต่างๆ

ขอกราบขอบพระคุณสมาชิกในครอบครัว ที่คอยช่วยเหลือและให้กำลังใจตลอดมา ขอขอบคุณนักศึกษาระดับบัณฑิตศึกษา ในสาขาวิชาวิศวกรรมอุตสาหกรรมและระบบ ที่ให้กำลังใจและคำปรึกษาจนทำให้วิทยานิพนธ์นี้เสร็จสมบูรณ์ด้วยดี และขอขอบคุณทุกๆ ท่านที่ไม่ได้กล่าวมา ณ ที่นี้ที่มีส่วนร่วมในการทำวิทยานิพนธ์ในครั้งนี้

นวพร ณะสาร



## สารบัญ

	หน้า
บทคัดย่อ	(5)
ABSTRACT	(6)
กิตติกรรมประกาศ	(7)
สารบัญ	(8)
สารบัญตาราง	(10)
สารบัญรูปภาพ	(11)
บทที่ 1 บทนำ	
1.1 ความสำคัญและที่มาของปัญหา	1
1.2 วัตถุประสงค์ของการวิจัย	5
1.3 ขอบเขตของงานวิจัย	5
1.4 ประโยชน์ที่คาดว่าจะได้รับ	6
บทที่ 2 ทฤษฎีและเอกสารที่เกี่ยวข้อง	
2.1 กุ้ง (shrimp)	7
2.2 กุ้งขาวลิโทพีเนียส แวนนาไม	8
2.3 ผลกระทบและกระบวนการแปรรูปกุ้ง	10
2.4 มาตรฐานกุ้งสดและแช่แข็งของประเทศสหรัฐอเมริกา	15
2.5 เส้นดำกุ้ง (vein)	17
2.6 ความรู้เบื้องต้นเกี่ยวกับการประมวลผลภาพดิจิทัล	18
2.7 การประมวลผลภาพ (image processing)	20
2.8 สถิติที่ใช้ในการเตรียมข้อมูล	30
2.9 เทคนิคการจำแนกกลุ่มในงานวิจัย	35
2.10 โปรแกรมที่ใช้ในงานวิจัย	43
2.11 การตรวจสอบเอกสาร	45

## สารบัญ

	หน้า
บทที่ 3 วิธีการดำเนินวิจัย	
3.1 ศึกษาข้อมูลและการวิเคราะห์เกณฑ์การตรวจสอบข้อบกพร่องของกึ่งแข็ง	51
3.2 การเตรียมตัวอย่าง และการออกแบบระบบดิงภาพ	51
3.3 การพัฒนาเทคนิคการประมวลผลภาพสำหรับตรวจสอบข้อบกพร่องในกึ่ง	57
3.4 การประเมินผล	75
บทที่ 4 ผลและการวิเคราะห์ผลการวิจัย	
4.1 ผลการวิเคราะห์เกณฑ์ในการจำแนก	77
4.2 ผลการเก็บข้อมูลภาพกึ่ง	79
4.3 ผลการพัฒนาเทคนิคการประมวลผลภาพสำหรับตรวจสอบคุณภาพของกึ่ง	79
4.4 ผลการประเมินผลประสิทธิภาพ	100
บทที่ 5 บทสรุปและข้อเสนอแนะ	
5.1 สรุปผลการดำเนินงานวิจัย	106
5.2 ข้อเสนอแนะในการดำเนินงานวิจัย	109
บรรณานุกรม	112
ภาคผนวก ก	119
ภาคผนวก ข	127
ภาคผนวก ค	137
ภาคผนวก ง	148
ประวัติผู้เขียน	156

## สารบัญตาราง

ตารางที่	หน้า
2.1	11
2.2	15
3.1	54
3.2	64
3.3	68
3.4	75
4.1	91
4.2	91
4.3	99
4.4	99
4.5	101
4.6	101

## สารบัญรูปภาพ

รูปที่		หน้า
1.1	มูลค่าการส่งออกผลิตภัณฑ์กุ้งของไทย (ล้านเหรียญสหรัฐฯ)	2
1.2	ตลาดส่งออกผลิตภัณฑ์กุ้งที่สำคัญปีพ.ศ.2559	2
1.3	ลักษณะบกพร่องที่ต้องตรวจสอบ 8 ประเภท	3
1.4	ลักษณะเส้นดำหลังกุ้ง	5
2.1	ลักษณะและส่วนประกอบของกุ้ง	8
2.2	กุ้งขาว หรือกุ้งขาวแวนนาไม	9
2.3	ขั้นตอนการผลิตกุ้งแช่แข็ง	13
2.4	เส้นดำหรือลำไส้ของกุ้ง	17
2.5	ตำแหน่งของพิกเซล	18
2.6	ภาพขาวดำ (binary image)	19
2.7	ภาพระดับเทา (gray scale image)	19
2.8	ภาพสี (color image)	20
2.9	ขั้นตอนพื้นฐานของการประมวลผลภาพด้วยคอมพิวเตอร์	21
2.10	ภาพกุ้ง	22
2.11	ความเข้มแสงเทียบกับค่าที่ได้จากเซลล์รับภาพ	23
2.12	การกำจัดสัญญาณรบกวน	25
2.13	การหาขอบภาพ	25
2.14	การแปลงภาพสีเป็นภาพระดับเทา	25
2.15	การแยกบริเวณโดยการใช้ค่าเทรชโฮลด์ (threshold)	27
2.16	การแยกบริเวณโดยการใช้ขอบของวัตถุ (edge base segmentation)	27
2.17	ฮีสโตแกรมภาพระดับเทา	28
2.18	การทำงานจริงที่ประกอบด้วยกลไกต่างๆ	29
2.19	ตำแหน่งข้อมูลสองกลุ่มในพีเจอาร์สเปซ	40
2.20	การวางตัวของข้อมูลในลักษณะเชิงเส้น	41
2.21	การเปลี่ยนมิติข้อมูลไปยังมิติที่สูงขึ้นด้วยฟังก์ชันของเคอร์เนล	42
2.22	โปรแกรม ImageJ เวอร์ชัน 1.51k	44
2.23	โปรแกรม Weka เวอร์ชัน 3.8.1	44

## สารบัญรูปร่างภาพ (ต่อ)

รูปที่		หน้า
2.24	มะเขือเทศสีไม่สม่ำเสมอใช้การแบ่งส่วนด้วยฮิสโตแกรม	47
2.25	การรับภาพ และแบ่งส่วนภาพของต้นกล้ามะเขือเทศ	47
2.26	การแบ่งส่วนและประมวลผลภาพกุ่ม	49
3.1	หัวข้อและวิธีดำเนินการวิจัย	50
3.2	กุ่มที่มีลักษณะเส้นดำตามเกณฑ์มาตรฐานการกำจัดเส้นดำ	52
3.3	ระบบดึงข้อมูลภาพ	53
3.4	กล้องบันทึกภาพดิจิทัล	54
3.5	แท่นถ่ายรูป ยี่ห้อ KAISER RSI	55
3.6	หลอดไฟฟลูออเรสเซนต์ชนิด Illuminant D (D65)	56
3.7	ขั้นตอนการพัฒนาเทคนิคการประมวลผลภาพสำหรับการตรวจสอบ	57
3.8	คำสั่งในโปรแกรม ImageJ สำหรับแปลงภาพสี RGB เป็นภาพระดับเทา	59
3.9	ชุดคำสั่งการแบ่งส่วนภาพด้วยการขีดแบ่งและการปรับค่าแบ่ง 2 ค่าจากฮิสโตแกรมของภาพ	60
3.10	คุณลักษณะเฉพาะที่สกัดได้จากภาพ	61
3.11	ชุดคำสั่งในการแยกคุณลักษณะทางรูปร่าง	62
3.12	การวัดคุณลักษณะทางรูปร่าง	62
3.13	กุ่มและฮิสโตแกรมของกุ่มที่ไม่ผ่านเกณฑ์มาตรฐาน	63
3.14	กุ่มและฮิสโตแกรมของกุ่มที่ผ่านเกณฑ์มาตรฐาน	63
3.15	กระบวนการจำแนกกลุ่มข้อบกพร่อง	65
3.16	วิธี cross-validation test	67
3.17	ผลการวิเคราะห์เบื้องต้นด้วยโปรแกรม WEKA	73
3.18	รายละเอียดของแอตทริบิวต์ที่เป็นลักษณะของจำนวนตัวเลข	73
3.19	ฮิสโตแกรมแสดงความถี่ของทุกแอตทริบิวต์เทียบกับคลาส (class) ปลายทาง	74
4.1	ลักษณะการแบ่งส่วนกุ่ม	78
4.2	ภาพตัวอย่างกุ่ม	80

### สารบัญรูปภาพ (ต่อ)

รูปที่		หน้า
4.3	ภาพการแยกบริเวณด้วยเทคนิคเทอร์สโฮล	81
4.4	ผลการหาความยาวของเส้นดำและตัวกึ่ง	82
4.5	ผลการหาพื้นที่ของเส้นดำและตัวกึ่ง	82
4.6	ผลการประมวลผลสีสโตแกรมภาพกึ่ง	83
4.7	รูปแบบการวิเคราะห์การจำแนกกลุ่มด้วย discriminant analysis	87
4.8	รูปแบบการวิเคราะห์การจำแนกกลุ่มด้วยซัพพอร์ตเวกเตอร์แมชชีน	92
4.9	ผลการสร้างโมเดลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนกรณีข้อมูลภาพไม่มีการลดขนาด	93
4.10	ผลการสร้างโมเดลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนกรณีข้อมูลภาพที่มีการลดขนาด	94
4.11	ผลการทดสอบโมเดลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนกรณีข้อมูลภาพไม่มีการลดขนาด	95
4.12	ผลการทดสอบโมเดลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนกรณีข้อมูลภาพมีการลดขนาด	95
4.13	ผลเปรียบเทียบประสิทธิภาพด้านความถูกต้อง (accuracy)	103
5.1	การประยุกต์เทคนิคประมวลผลภาพกับเครื่องจักรอัตโนมัติ	110

# บทที่ 1

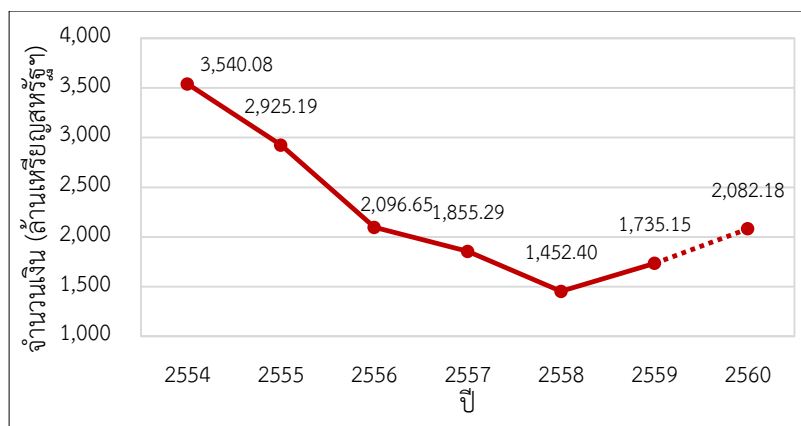
## บทนำ

### 1.1 บทนำต้นเรื่อง

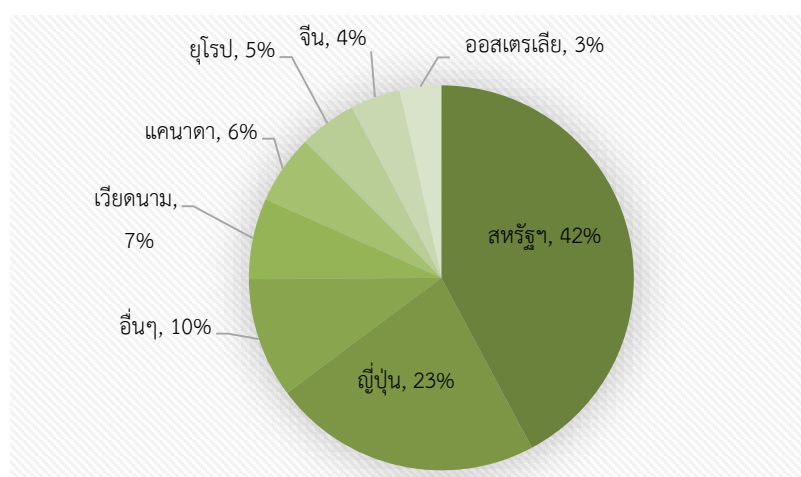
อุตสาหกรรมอาหารเป็นอุตสาหกรรมที่สร้างรายได้จากการส่งออกให้แก่ประเทศไทย เป็นมูลค่าสูง และมีส่วนสำคัญอย่างยิ่งต่อโครงสร้างระบบเศรษฐกิจของประเทศ ดังจะเห็นได้ว่าในปี 2559 อุตสาหกรรมประเภนี้ก่อให้เกิดรายได้เป็นมูลค่ามากกว่า 950,114 ล้านบาท ซึ่งถือเป็นอันดับ 1 ของอาเซียน และคาดว่าจะมีแนวโน้มขยายตัวเพิ่มขึ้นเป็น 1,050,000 ล้านบาทในปี 2560[1] ปัจจุบันอุตสาหกรรมอาหารไทยสามารถส่งออกสินค้าไปยังประเทศต่าง ๆ ทั่วโลกได้มากกว่า 200 ประเทศ[2] ด้วยความอุดมสมบูรณ์ทางด้านวัตถุดิบทั้งจากภาคเกษตรกรรม ประมงและปศุสัตว์ ประเทศไทยนับว่ามีศักยภาพสูงในด้านการผลิตและส่งออกอาหาร ทั้งในรูปของวัตถุดิบและผลิตภัณฑ์ ที่ผ่านการแปรรูปจากวัตถุดิบแล้ว

ผลิตภัณฑ์กุ้งแช่แข็งเป็นผลิตภัณฑ์ในอุตสาหกรรมอาหารจากการประมงประเภทหนึ่งที่มีความสำคัญของประเทศไทยและมีมูลค่าการส่งออกเป็นอันดับ 1 ของโลก แต่ในปี 2559 มูลค่าการส่งออกได้ลดต่ำลงเป็นอันดับ 3 รองจากประเทศเอกวาดอร์และอินเดีย โดยมีมูลค่าส่งออกเป็นปริมาณ 1,735.15 ล้านเหรียญสหรัฐฯ[3] ดังรูปที่ 1.1 ปริมาณการส่งออกที่ลดลงในช่วงปี 2558 นี้เป็นผลสืบเนื่องมาจากโรคระบาดและภัยแล้ง อย่างไรก็ตามในปี 2560 ผลผลิตกุ้งมีแนวโน้มเพิ่มขึ้น เนื่องจากสามารถแก้ปัญหาดังกล่าวได้อีกทั้งประเทศคู่ค้าต่างเชื่อมั่นในคุณภาพของผลิตภัณฑ์กุ้งไทย จึงคาดว่าประเทศไทยจะกลับมาเป็นผู้ส่งออกกุ้งอันดับที่ 1 ได้อีกครั้ง ปัจจุบันตลาดส่งออกสำคัญของประเทศไทย ได้แก่ สหรัฐอเมริกา ญี่ปุ่นและเวียดนาม[4] ซึ่งมีสัดส่วนดังแสดงในรูปที่ 1.2

กุ้งสดเป็นวัตถุดิบหลักในการผลิตกุ้งแช่แข็ง ได้ถูกนำเข้าสู่กระบวนการผลิตและแปรรูปผ่านขั้นตอนต่างๆ ตั้งแต่การคัดขนาด ชั่งน้ำหนัก ทำความสะอาด ตัดแต่ง เช่น หักหัว ปอกเปลือก ผ่าหลัง เป็นต้น หลังจากนั้นจะผ่านกระบวนการแปรรูปผลิตภัณฑ์ตามขั้นตอนเพิ่มเติมของแต่ละผลิตภัณฑ์ เช่น การยืด การต้ม การปรุงรส การชุบเกล็ดขนมปัง เป็นต้น สุดท้ายผลิตภัณฑ์จะถูกบรรจุและนำไปแช่ในอุณหภูมิที่เหมาะสม[5] ทั้งนี้คุณภาพของผลิตภัณฑ์เป็นสิ่งสำคัญที่สร้างความเชื่อมั่นให้แก่ผู้บริโภค นอกจากความสด ความสะอาด และรสชาติแล้ว ความสวยงาม และความปลอดภัยยังเป็นสิ่งที่ต้องคำนึงถึงอีกด้วย ปัจจุบันประเทศคู่ค้าที่นำเข้าผลิตภัณฑ์กุ้งแช่แข็งได้มีการกำหนดประเภทของลักษณะข้อบกพร่องที่ต้องตรวจสอบ 8 ลักษณะ[6] ประกอบด้วย



รูปที่ 1.1 มูลค่าการส่งออกผลิตภัณฑ์กุ้งของไทย (ล้านเหรียญสหรัฐฯ)[3]



รูปที่ 1.2 ตลาดส่งออกผลิตภัณฑ์กุ้งที่สำคัญปีพ.ศ.2559[4]

- 1) การสูญเสียน้ำของเนื้อกุ้ง (รูปที่ 1.3 (ก)) สามารถตรวจสอบโดยการสังเกตการคายน้ำจากบริเวณผิวของกุ้งว่ามีลักษณะเหนียว แห้ง มีสีขาวหรือเหลืองปกคลุม
- 2) ความไม่สม่ำเสมอของขนาด (รูปที่ 1.3 (ข)) สามารถตรวจสอบได้จากการสุ่มตัวอย่างกุ้งมาทำการวัดขนาด และชั่งน้ำหนัก
- 3) ลักษณะจุดดำดำบนตัวกุ้ง (รูปที่ 1.3 (ค)) สามารถตรวจสอบได้จากการสังเกตการปรากฏของจุดที่ขนาดใหญ่กว่าปลายปากกาหรือพื้นที่สีดำที่สามารถมองเห็นได้
- 4) ความไม่สมบูรณ์ของตัวกุ้ง (รูปที่ 1.3 (ง)) สามารถตรวจสอบได้จากการนับและการสังเกตจำนวนข้อตามขนาดของกุ้งว่ามี การหัก ขาด หรือเสียหายของลำตัวกุ้ง
- 5) การเจือปนของชิ้นส่วนที่ไม่ต้องการ (รูปที่ 1.3 (จ)) สามารถตรวจสอบได้จากการสังเกตวัสดุเจือปนต่างๆ เช่น ขา หนวด ครีบ หรือวัสดุภายนอกอื่นๆที่ไม่ใช่ชิ้นส่วนของกุ้งแต่ไม่อันตรายต่อผู้บริโภค





โดยปกติแล้ว ในอุตสาหกรรมอาหารทะเลแปรรูปมีการใช้แรงงานคนเป็นหลักในการทำงาน (labor intensive)[10] โดยเฉพาะอย่างยิ่งในกระบวนการตรวจสอบคุณภาพที่ต้องอาศัยทักษะและประสบการณ์ ผู้ปฏิบัติงานที่ทำหน้าที่ในกระบวนการนี้ต้องผ่านการอบรมเพื่อให้สามารถตรวจสอบผลิตภัณฑ์ได้ตามเกณฑ์ที่กำหนดไว้[11] ทำให้ในปัจจุบันเกิดปัญหาที่เกี่ยวข้องกับแรงงานหลายกรณี เช่น การขาดแคลนแรงงาน แรงงานมีอัตราการหมุนเวียนเข้าออกสูง และประสิทธิภาพในการทำงานของแรงงานต่ำ[12] จากการสำรวจพบว่าในการตรวจสอบคุณภาพด้วยผู้ปฏิบัติงานมีความถูกต้องประมาณ 60%[13], [14] เนื่องจากในขั้นตอนนี้ผู้ปฏิบัติงานต้องใช้สายตาและประสาทสัมผัสในการประเมินเพื่อแยกแยะและตัดสินใจ โดยประสิทธิภาพการทำงานของผู้ปฏิบัติงานจะค่อยๆลดลงตามระยะเวลาที่ผ่านมา ซึ่งอาจเกิดจากความเมื่อยล้าและข้อจำกัดอื่นๆของร่างกาย[11] ทำให้ผู้ผลิตต้องมีค่าใช้จ่ายเพิ่มเติมในการอบรมและฝึกทักษะพนักงานใหม่ และคาดว่าในอีก 5 ปีข้างหน้าปัญหานี้มีแนวโน้มความรุนแรงเพิ่มมากขึ้น[15] โดยขณะนี้ประเทศไทยมีการกำหนดยุทธศาสตร์เพื่อการขับเคลื่อนอุตสาหกรรม 4.0 ด้วยการส่งเสริมและสนับสนุนภาคอุตสาหกรรมให้เข้าถึงและสามารถใช้ระบบการผลิตแบบอัตโนมัติ และเทคโนโลยีสารสนเทศ[16] ซึ่งปัจจุบันนี้จึงพบว่าผู้ผลิตหลายรายได้มีการนำเทคโนโลยีสมัยใหม่มาใช้ในกระบวนการผลิตและตรวจสอบ[17] เช่น เครื่องทำความสะอาดเครื่องคัดขนาด เครื่องปอกเปลือก ผ่าหลัง และกำจัดเส้นดำ เป็นต้น เพื่อควบคุมต้นทุนการผลิต ควบคุมมาตรฐานผลิตภัณฑ์และตรวจสอบคุณภาพของผลิตภัณฑ์ ซึ่งสามารถช่วยลดผลกระทบจากปัญหาแรงงานดังกล่าวได้

จากการสำรวจ พบว่าเทคโนโลยีหนึ่งที่มีบทบาทเป็นอย่างมากในกระบวนการตรวจสอบคุณภาพอัตโนมัติคือ เทคโนโลยีการประมวลผล แม้ว่าปัจจุบันมีผู้วิจัยด้านประมวลผลภาพในการตรวจสอบลักษณะบกพร่องในผลิตภัณฑ์กันอย่างกว้างขวาง เช่น การพัฒนาเทคนิคเพื่อใช้ในการประมาณค่าความชื้นในกุ้ง[18] เทคนิคประมวลผลด้วยภาพในการตรวจสอบเปลือกกุ้งสด[19] เทคนิคประมวลผลด้วยภาพการตรวจสอบจุดดำบนตัวกุ้ง[20] เทคนิคประมวลผลด้วยภาพเพื่อหาขนาดและน้ำหนัก[21]–[23] เทคนิคประมวลผลด้วยภาพเพื่อตรวจสอบความสมบูรณ์ของกุ้ง[8], [24], [25] และเทคนิคประมวลผลด้วยภาพเพื่อตรวจสอบสีของกุ้ง[26] เป็นต้น แต่ยังไม่มีการวิจัยใดที่เกี่ยวข้องการใกล้เคียงกับการนำเทคนิคการประมวลผลภาพเพื่อใช้ในกระบวนการตรวจสอบเส้นดำที่กำจัดไม่หมด ดังที่กล่าวแล้วว่าเส้นดำที่หลงเหลืออยู่บนลำตัวกุ้งเป็นลักษณะบกพร่องชนิดหนึ่งในผลิตภัณฑ์ที่มีการกำจัดเส้นดำ ดังรูปที่ 1.4 ซึ่งเส้นดำ (vein หรือ gut) เป็นส่วนหนึ่งของระบบย่อยอาหารของกุ้งจะมีสีและความเข้มแตกต่างกันขึ้นกับพฤติกรรมการกินอาหาร โดยทั่วไปจะมีสีดำและจะมีความเข้มขึ้นจากการกินพืชหรือสัตว์อื่นๆในบ่อ[27] ถึงแม้ว่าเส้นดำที่กำจัดไม่หมดจะไม่ส่งผลกระทบต่อรสชาติ และไม่เป็นอันตรายต่อผู้บริโภค แต่ในผู้ที่ไวต่อความสะอาดเส้นดำที่หลงเหลืออยู่อาจทำให้มีผลต่อสุขภาพได้ เช่น อาการท้องเสีย ปวดท้อง เป็นต้น เพราะภายในเส้นดำมีเศษอาหาร กวาด

และสิ่งสกปรกต่าง ๆ [6] อีกทั้งความยาวของเส้นด้ายได้ถูกควบคุมเป็นข้อกำหนดหนึ่งของเกณฑ์การตรวจสอบคุณภาพของผลิตภัณฑ์กึ่งแข็งในชนิดที่ต้องมีการกำจัดเส้นด้าย ซึ่งการตรวจสอบเส้นด้ายสามารถทำได้จากการสังเกตจากสายตาของผู้ปฏิบัติงานเพียงเท่านั้น



รูปที่ 1.4 ลักษณะเส้นด้ายหลังกึ่ง

งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อพัฒนาระบบอัตโนมัติที่ใช้ในการคัดแยกกึ่งแข็งที่กำจัดเส้นด้ายไม่หมดด้วยการประมวลผลภาพ ซึ่งเมื่อผลการวิจัยสำเร็จเทคนิคที่ได้พัฒนาขึ้นนี้จะสามารถนำไปประยุกต์ใช้กับเครื่องมือระบบอัตโนมัติที่มีในปัจจุบันร่วมกับการตรวจสอบคุณภาพด้วยเทคนิคประมวลผลภาพอื่นๆ ซึ่งทำให้กระบวนการทำงานเป็นระบบอัตโนมัติที่ครบวงจรสามารถผลิตและตรวจสอบข้อบกพร่องของกึ่งได้อย่างมีประสิทธิภาพทดแทนกระบวนการทำงานด้วยผู้ปฏิบัติงานได้ เพื่อความสม่ำเสมอและความถูกต้องในการปฏิบัติงาน อีกทั้งสามารถนำไปประยุกต์ใช้กับงานด้านการตรวจสอบคุณภาพกับสินค้าเกษตรอื่นๆ ได้

## 1.2 วัตถุประสงค์ของงานวิจัย

เพื่อพัฒนาเทคนิคสำหรับตรวจสอบข้อบกพร่องเฉพาะเส้นด้ายกลางหลังของกึ่ง โดยประยุกต์ใช้เทคนิคประมวลผลภาพ และการจำแนกกลุ่มซึ่งจะมีความถูกต้องในการตรวจสอบอย่างน้อย 95% เมื่อเทียบกับลักษณะตามมาตรฐานของผลิตภัณฑ์นำเข้าของประเทศคู่ค้า

## 1.3 ขอบเขตของการวิจัย

1) เครื่องมือและอุปกรณ์ที่สร้างขึ้นจะยังไม่เป็นระบบอัตโนมัติ เป็นเพียงระบบเพื่อทำการทดลองประเมินเปรียบเทียบความถูกต้องของเทคนิคที่จะถูกพัฒนาขึ้นกับเกณฑ์มาตรฐานการนำเข้ากึ่งแข็งของประเทศสหรัฐอเมริกา

2) เทคนิคการตรวจสอบคุณภาพในการวิจัยนี้ จะเป็นการตรวจสอบคุณภาพภายนอกของผลิตภัณฑ์กึ่งขาลิโทพีเนียสแวนนาไมสขนาดน้อยกว่า 70 ตัวต่อปอนด์ที่มีการผ่าหลังเพื่อกำจัดเส้นด้ายออก

3) การจัดวางกึ่ง (scene constraint) เพื่อรับภาพนำไปประมวลผลกึ่งมีลักษณะวางนอนตามยาว และหันส่วนหลังขึ้น เนื่องจากกระบวนการตรวจสอบนี้จะนำมาใช้หลังกระบวนการตัดแต่งวัตถุดิบ

#### 1.4 ประโยชน์ที่คาดว่าจะได้รับ

1) เมื่อนำไปใช้ในภาคอุตสาหกรรมเทคนิคนี้จะสามารถตรวจสอบคัดแยกกึ่งที่มีการกำจัดเส้นด้ายไม่ถูกต้องได้อย่างแม่นยำทดแทนกระบวนการตรวจสอบคุณภาพด้วยผู้ปฏิบัติงานได้

2) สามารถนำเทคนิคการประมวลผลภาพในส่วนของวัสดุอุปกรณ์ ขั้นตอนของกระบวนการ และวิธีการจำแนกกลุ่มไปประยุกต์ใช้กับการตรวจสอบและคัดแยกผลิตภัณฑ์อื่นๆ โดยปรับเปลี่ยนจำนวนพารามิเตอร์ และคุณลักษณะเฉพาะของวัตถุดิบนั้น รวมทั้งการแบ่งข้อมูลเพื่อสอนและทดสอบโมเดล การเตรียมข้อมูลก่อนนำไปวิเคราะห์ด้วยเทคนิคต่างๆ และวิธีการจำแนกให้เหมาะสมกับลักษณะของงานวิจัย

## บทที่ 2

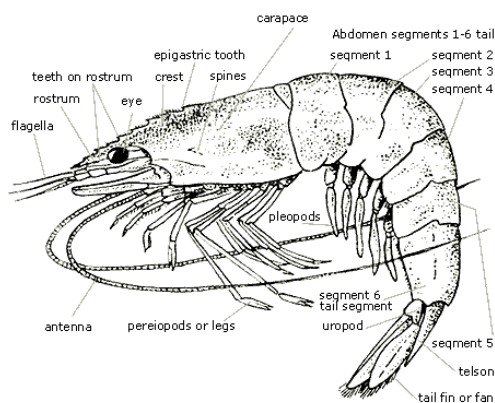
### ทฤษฎีและวิจัยที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงทฤษฎีที่เกี่ยวข้องในงานวิจัยการพัฒนาระบบประมวลผลภาพการตรวจสอบข้อบกพร่องในกุ้ง ได้แก่ กุ้ง[27]–[30] กระบวนการแปรรูปผลิตภัณฑ์กุ้ง มาตรฐานการตรวจสอบคุณภาพกุ้ง ความรู้พื้นฐานของกระบวนการประมวลผลภาพ และเทคนิคที่ใช้จำแนกระดับคุณภาพรวมถึงงานวิจัยที่เกี่ยวข้องซึ่งจะกล่าวรายละเอียดในหัวข้อถัดไป

#### 2.1 กุ้ง (shrimp) ,

กุ้งเป็นสัตว์น้ำในกลุ่ม crustacean อยู่ในกลุ่มเดียวกับกุ้ง และปู เป็นสัตว์ที่ไม่มีกระดูกสันหลังเป็นพวกที่ว่ายน้ำ จัดอยู่ในไฟลัมอาร์โทรพอดา (Arthropoda) เป็นอาร์โทรพอด (Arthropod) หายใจด้วยเหงือก กุ้งเป็นสัตว์เลือดเย็น มี 10 ขา หรือ 5 คู่ มีลำตัวเป็นปล้อง โดยมีรยางค์ยื่นออกมาเป็นคู่ เช่น หนวด ขากรรไกร ขาเดิน และขาว่ายน้ำลำตัวแบนข้างขาว่ายน้ำ (Pleopods) เป็นรยางค์ส่วนของลำตัวที่พัฒนาและทำหน้าที่ว่ายน้ำโดยตรง ส่วนท้อง (Thoracic Sternum) ซึ่งอยู่ระหว่างขาว่ายน้ำทั้ง 2 ข้าง แคบ และสังเกตเห็นได้ยาก เปลือกด้านข้างตัว (Lateral Pleuron) แบ่งเป็นแผ่นชัดเจนในแต่ละปล้อง และส่วนปลายหาง (Telson) จะเรียวยาวทางด้านท้ายเหมือนหนามในกุ้งระยะหลังวัยอ่อนจนโตเต็มวัยดังรูปที่ 2.1 เมื่อเจริญเติบโตจะสลัดเปลือกเดิม สร้างเปลือกใหม่ ซึ่งเปลือกหุ้มตัวเป็นสารกลุ่มไคติน โดยปกติชอบหลบซ่อนตัวอยู่เงียบๆ ตามพื้นน้ำ จะออกหากินในเวลากลางคืน กุ้งกินทั้งพืชและสัตว์เป็นอาหาร เช่น กินกุ้งด้วยกันเอง ลูกปลา ไส้เดือน และสัตว์หน้าดินขนาดเล็กชนิดต่าง ๆ[28]

กุ้งเป็นอาหารทะเลที่มีเนื้ออร่อย รับประทานกันอย่างแพร่หลายทั่วโลก ลักษณะที่ต้องการของกุ้งสด คือ มี รูปร่างปกติและมีอวัยวะครบ มีสีตามธรรมชาติ ส่วนหัวและลำตัวติดแน่น เปลือกสดใส ลื่น เป็นเงามันตามธรรมชาติ เนื้อแน่น ใสติดกับเปลือกแน่นแกะออกจากเปลือกยากกุ้ง [29]



รูปที่ 2.1 ลักษณะและส่วนประกอบของกุ้ง[31]

กุ้งที่รู้จักกันดีในปัจจุบันนี้มีประมาณ 3,047 ชนิด แบ่งออกเป็น 4 กลุ่มใหญ่ ๆ คือ

1) Sargestoidea (ประมาณ 94 ชนิด) มีลักษณะทั่วไปคล้ายกุ้ง ลำตัวบาง ขนาดเล็ก สำหรับประเทศไทยใช้ทำเป็นกะปิ เรียกกุ้งเคย มีหลายชนิด เช่น เคยสารส้มโอ เคยตาดำ เป็นต้น

2) Penaeoidea (ประมาณ 376 ชนิด) กุ้งในกลุ่มนี้ส่วนมากเป็นกุ้งทะเล มีหลายชนิดที่มีการเลี้ยง และผลผลิตสูง ทำให้มีความสำคัญทางเศรษฐกิจมาก

3) Stenopoidea (ประมาณ 60 ชนิด) กุ้งในกลุ่มนี้มีขนาดเล็ก พบเฉพาะในทะเล มีความสำคัญทางเศรษฐกิจน้อย และจับได้ปริมาณไม่มาก มีวงศ์ (family) เดียว มักจะพบตามแนวปะการัง และอาจจะถูกจับมาขายเป็นสัตว์เลี้ยงในตู้กระจก จึงมีคุณค่าทางเศรษฐกิจบ้าง

4) Caridea (ประมาณ 2,517 ชนิด) อาศัยได้ทั้งในน้ำจืด น้ำกร่อย และบางชนิดพบเฉพาะในทะเล

แม้ว่ากุ้งในกลุ่มที่ 4 มีจำนวนชนิดมากที่สุด แต่มีความสำคัญทางเศรษฐกิจไม่สูงเท่ากับกลุ่ม Penaeoidea กุ้งที่มีความสำคัญทางเศรษฐกิจที่จับได้ในปัจจุบันมีประมาณ 300 ชนิด และเพียง 100 ชนิดเท่านั้นที่เคยปรากฏในรายงานสถิติการทำประมงของโลก โดยในงานวิจัยนี้ศึกษา กุ้งในกลุ่ม Penaeoidea ที่นิยมนำมาผลิตกุ้งแช่แข็งในปัจจุบัน คือ กุ้งขาวแวนนาไม

## 2.2 กุ้งขาวลิโทพีเนียสแวนนาไม

### 1) ลักษณะทั่วไป

กุ้งขาวลิโทพีเนียสแวนนาไม มีลำตัวสีขาว 8 ปล้อง หน้าอกใหญ่การเคลื่อนไหวเร็ว ส่วนหัวมี 1 ปล้อง มีกรืออยู่ในระดับยาวประมาณ 0.8 เท่าของความยาวเปลือกหัวสั้นกริสสูง ปลาย

กรีกแคบ ส่วนของกรีกมีลักษณะเป็นสามเหลี่ยมมีสีแดงอมน้ำตาล กรีกด้านบนมี 8 ฟัน กรีกด้านล่างมี 2 ฟัน ร่องบนกรีกมองเห็นได้ชัด เปลือกหัวสีขาวอมชมพูถึงแดง ขาเดินมีสีขาวเป็นลักษณะที่โดดเด่น หนวดแดง 2 เส้นยาว ตาแดงเข้ม ส่วนตัวมี 6 ปล้อง เปลือกตัวสีขาวอมชมพูถึงแดง เปลือกบาง ขาวายน้ำ 5 คู่ มีสีขาวข้างในที่ปลายมีสีแดง ส่วนหางมี 1 ปล้อง ปลายหางมีสีแดงเข้ม แพนหางมี 4 ใบและ 1 กรีก หาง ขนาดตัวที่โตสมบูรณ์เต็มที่ของกุ้งสายพันธุ์นี้จะมีขนาดเล็กกว่ากุ้งกุลาล่าดำ หากินทุกระดับความลึกของน้ำ ชอบว่ายน้ำล่องน้ำแก่ง ลอกคราบเร็วทุกๆสัปดาห์ ไม่หมกตัว[27] ดังรูปที่ 2.2



รูปที่ 2.2 กุ้งขาว หรือกุ้งขาวแวนนาไม[32]

กุ้งขาวลิโทพีเนียสแวนนาไมเป็นสายพันธุ์กุ้งทะเลที่มีการเพาะเลี้ยงกันอย่างแพร่หลายในหลายประเทศ เช่นสหรัฐอเมริกา เม็กซิโก กัวเตมาลา นิคารากัว คอสตาริกา ปานามา โคลัมเบีย อีควาดอร์ เปรู กุ้งสายพันธุ์นี้เป็นสัตว์ที่มีความแข็งแรงและทนทานจึงมีการขยายพันธุ์ตามธรรมชาติได้กว้างไกล ในแถบแนวชายฝั่งตะวันออกของมหาสมุทรแปซิฟิก ตั้งแต่เม็กซิโกถึงเปรู เนื่องจากภูมิภาคในแถบนี้ที่ระดับความลึกจากเส้นแนวชายฝั่งลงไปประมาณ 72 เมตรหรือ 235 ฟุต มีพื้นที่ท้องทะเลเป็นเหมือนโคลนที่เหมาะสมแก่การเจริญเติบโต และเป็นแหล่งอาหารที่อุดมสมบูรณ์ ประเทศอีควาดอร์เป็นประเทศผู้ผลิตรายใหญ่ที่มีฟาร์มเพาะเลี้ยงกุ้งลูกกุ้ง พ่อ-แม่พันธุ์

## 2) วงจรชีวิต

ในธรรมชาติกุ้งขาวลิโทพีเนียสแวนนาไมจะมีอายุประมาณเกือบ 36 เดือน โดยจะวางไข่ที่ระดับน้ำลึกประมาณ 30-60 มิลลิเมตรใกล้พื้นทรายปกติแล้วแม่กุ้งขนาด 60-120 กรัม จะวางไข่ประมาณ 150,000 ถึง 250,000 ฟอง ส่วนแม่กุ้งขนาด 30-45 กรัม จะวางไข่ประมาณไม่เกิน 100,000 ฟอง โดยจะวางไข่ในตอนกลางคืนบนพื้น แม่กุ้งจะว่ายน้ำอย่างรวดเร็วอยู่ประมาณ 45-60

วินาที แล้วจึงเริ่มออกไข่ขณะที่ลดความเร็วลงอย่างช้าๆ เนื่องจากลักษณะอวัยวะสืบพันธุ์เพศเมียของ กุ้งขาวลิโทพีเนียสแวนนาไม่นี้จะมีลักษณะเป็นแบบปิด (opened thelycum)

### 3) สภาพแวดล้อมในการเลี้ยง

กุ้งขาวลิโทพีเนียสแวนนาไมเป็นกุ้งที่เลี้ยงได้ทั้งระบบธรรมชาติ และระบบกึ่ง หนาแน่น ลักษณะพิเศษของกุ้งสายพันธุ์นี้คือสามารถสร้างความคุ้นเคยหรือปรับลักษณะนิสัยภายใต้ ระบบการเพาะเลี้ยงได้ สามารถเจริญเติบโตได้ในน้ำที่มีระดับความเค็มที่ 10-22 ส่วนในพันส่วนที่ อุณหภูมิ 26-29 องศาเซลเซียส ระดับออกซิเจนที่ละลายในน้ำควรมีค่า 4-9 มิลลิกรัมต่อลิตร และ สำหรับค่าความเป็นกรดและด่างควรอยู่ระหว่าง 7.2-8.6 ซึ่งสามารถทำการเพาะเลี้ยงได้ทั้งในบริเวณ พื้นที่ชายฝั่ง หรือบริเวณพื้นที่ที่มีความเค็มต่ำ กุ้งชนิดนี้ชอบน้ำกระด้างที่มีความกระด้างรวม 120 มิลลิกรัมต่อลิตร มีค่าอัลคาไลน์ในช่วง 80-150 มิลลิกรัมต่อลิตร มีนิสัยที่ไวต่อการเปลี่ยนแปลงสภาวะ ของน้ำในบ่อเพาะเลี้ยง ตื่นตกใจง่าย

### 4) อาหาร

อาหารที่ใช้เลี้ยงกุ้งชนิดนี้เป็นอาหารที่เกิดขึ้นในบ่อจากการที่หว่านอาหาร ชีวภาพสำหรับสัตว์หน้าดินต่างๆ อาหารอัดเม็ด อาหารเสริมแร่ธาตุในรูปคีเลต อาหารธรรมชาติ เช่น สาหร่าย แครอท ฟักทอง กัลฉ่าย มะเขือเทศ เพื่อเพิ่มเอนไซม์ และปริมาณสารแอสตาแซนทิน (astaxantine)

## 2.3 ผลิตภัณฑ์และกระบวนการแปรรูปกุ้ง

### 1) ผลิตภัณฑ์กุ้งสดแช่แข็ง


สำหรับรูปแบบผลิตภัณฑ์กุ้งสดแช่แข็งที่นิยมผลิตและเป็นที่ต้องการของตลาด แบ่ง ได้ทั้งหมด 5 แบบ แสดงดังตารางที่ 2.1 รูปแบบของผลิตภัณฑ์จากกุ้ง[33], [34]



ตารางที่ 2.1 รูปแบบของผลิตภัณฑ์จากกุ้งสด[35]

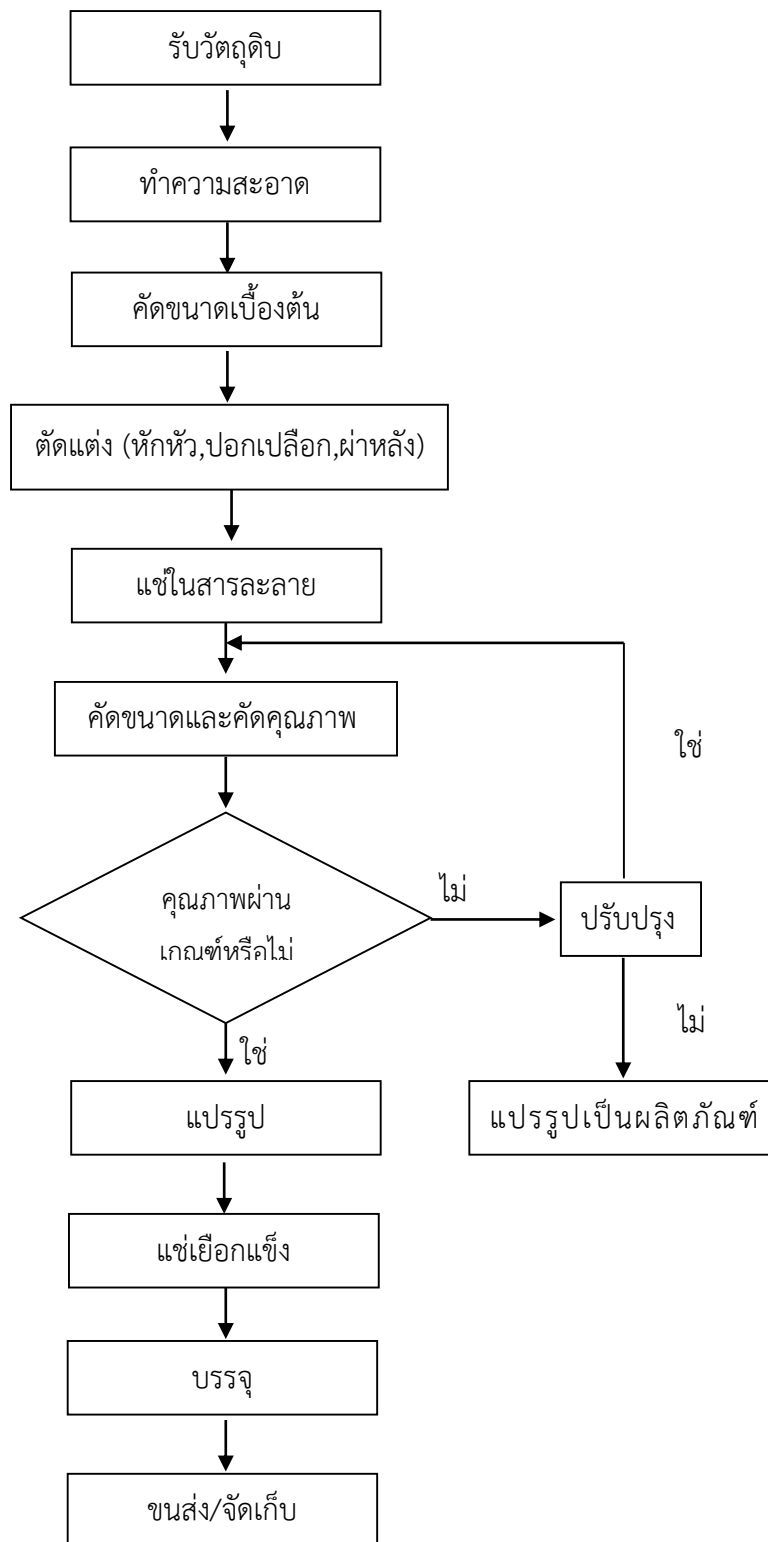
ผลิตภัณฑ์	รูปแบบ	ภาพประกอบ
<p>1. กุ้งสดทั้งตัว (Head-on, Shell-on)</p>	<p>กุ้งที่อยู่ในสภาพธรรมชาติที่มีหัวเปลือก และหางอยู่ครบ</p>	
<p>2. กุ้งขาวเด็ดหัวไว้หาง (Headless, Shell-on)</p>	<p>กุ้งที่เด็ดหัวออก แต่ลำตัวยังมีเปลือกและหางครบ มีทั้งชนิดซักไส้และไม่ซักไส้</p>	
<p>3. กุ้งสดปอกเปลือกและไว้หาง (Peeled, Tail-on)</p>	<p>กุ้งที่เอาหัวและเปลือกออก แต่ยังคงเหลือเปลือกข้อสุดท้ายติดกับแพนหาง มีทั้งชนิดซักไส้และไม่ซักไส้</p>	
<p>4. กุ้งสดปอกเปลือกไม่ไว้หาง (Peeled, Tail-off)</p>	<p>กุ้งที่เอาหัวและเปลือกออกทั้งตัว มีทั้งชนิดซักไส้และไม่ซักไส้</p>	

ตารางที่ 1.1 รูปแบบของผลิตภัณฑ์จากกุ้งสด (ต่อ) [35]

ผลิตภัณฑ์	รูปแบบ	ภาพประกอบ
5. กุ้งยืดหรือกุ้ง โนบาชิ (Strength shrimp)	กุ้งลอกเปลือก หางออก และผ่าหลัง เอาไส้ แล้วนำไปบั้งนำมายืดให้ตรงโดย อุปกรณ์ยืด	

## 2) กระบวนการผลิต

อาหารทะเลแช่แข็งเป็นผลิตภัณฑ์ที่วัตถุดิบต้องผ่านกระบวนการต่างๆตามประเภทของผลิตภัณฑ์ ได้แก่ การเตรียมเบื้องต้น (raw material preparation) เช่น คัดขนาด การล้าง แล่เป็นชิ้น ตัดแต่ง แกะเปลือก เป็นต้น การเพิ่มมูลค่า เช่น การชุบด้วยแป้ง การชุบเกล็ดขนมปัง เป็นต้น การปรุงสุก (cooking) เช่น นึ่ง ทอด ต้มสุก เป็นต้น การแปรรูปหรือนำไปประกอบอาหาร[5] โดยมีแผนผังกระบวนการผลิตรูปที่ 2.3



รูปที่ 2.3 ขั้นตอนการผลิตแกงแค[5]

ขั้นตอนการแปรรูปอาหารทะเลแช่เยือกแข็ง มีรายละเอียดของขั้นตอนดังนี้

(1) การรับวัตถุดิบ การเก็บรักษาคุณภาพอาหารทะเลก่อนส่งโรงงานแปรรูป ผู้จำหน่ายจะควบคุมอุณหภูมิของอาหารทะเลให้ต่ำกว่า 4 องศาเซลเซียส การลดอุณหภูมิทำโดยการใช้ น้ำแข็งผสมเกลือ หรือการเก็บรักษาไว้ในน้ำทะเลผสมน้ำแข็งที่เย็นจัดหรือเก็บรักษาในห้องเย็น (cold storage) วัตถุประสงค์ของการลดอุณหภูมิเพื่อป้องกันการเสื่อมเสียด้วยจุลินทรีย์ (microbial spoilage) เช่น Proteus และป้องกันการเพิ่มจำนวนของแบคทีเรียก่อโรค (pathogen) และรักษา ความสดของวัตถุดิบการตรวจสอบคุณภาพเพื่อรับวัตถุดิบ

(2) การล้าง (washing) เพื่อกำจัดสิ่งปนเปื้อน และลดปริมาณจุลินทรีย์เริ่มต้น ในวัตถุดิบ การล้างทำความสะอาดวัตถุดิบอาจใช้น้ำสะอาดเย็นผสมคลอรีนที่ความเข้มข้นระดับ ประมาณ 3-5% และอาจเติมเกลือเพื่อลดอุณหภูมิของน้ำ การล้างวัตถุดิบอาจกระทำโดยการใช้วิธีจุ่ม ล้างในภาชนะ หรือล้างผ่านสายพานที่มีน้ำฉีดล้างบนสายพานนั้น หรือทั้งสองวิธีร่วมกับการล้างใน ขั้นตอนแรกน้ำที่ผ่านการล้างจะมีการเจือปนด้วยเลือดและสิ่งเจือปนอื่นๆ ที่ติดมากับวัตถุดิบเป็น ปริมาณสูง ดังนั้นจึงต้องมีการล้างเพิ่มอีกหลายครั้งจนกระทั่งปริมาณจุลินทรีย์ที่ติดมากับวัตถุดิบมีค่าต่ำ กว่ามาตรฐาน

(3) การตัดแต่งขั้นต้น กุ้งจะถูกถอดหัว แกะเปลือก ผ่าหลัง ตามลักษณะของ ผลิตภัณฑ์ที่ต้องการ

(4) การแช่ในสารละลาย (soaking) เพื่อปรับปรุงเนื้อสัมผัส ช่วยให้อุ้มน้ำได้ดี โดยการแช่ในสารอุ้มน้ำ (moisture/water retention agent)

(5) การตัดแต่งขั้นสุดท้าย วัตถุดิบที่ผ่านขั้นตอนการตัดแต่งขั้นต้น ซึ่งมีการ กำจัดส่วนที่ไม่ต้องการออกไปแล้ว จะถูกตัดแต่งให้มีลักษณะตามต้องการของลูกค้า หรือตามที่ผู้ผลิต เองกำหนด จากนั้นวัตถุดิบจะถูกล้างทำความสะอาดอีกครั้งด้วยน้ำเย็น เพื่อกำจัดเศษเนื้อจากการตัด แต่งออกให้หมด ระหว่างการตัดแต่ง อาจมีการใช้น้ำแข็งรักษาอุณหภูมิ เพื่อคงคุณภาพความสดและ ลดอัตราการเจริญของจุลินทรีย์

(6) การคัดขนาด การคัดเกรด วัตถุดิบ จะถูกนำไปทำให้สะเด็ดน้ำก่อนทำการ คัดขนาด คุณภาพ และชั่งน้ำหนัก (weight) เพื่อบรรจุลงบรรจุภัณฑ์ต่อไป

(7) การแช่เยือกแข็ง รูปแบบโดยทั่วไปของอาหารทะเลแช่เยือกแข็ง ได้แก่

- อาหารทะเลแช่เยือกแข็งแบบเป็นชิ้น หรือที่เรียกว่าไอ คิว เอฟ (IQF) เป็นอาหารทะเลที่ภายหลังการแช่เยือกแข็ง จะมีลักษณะเป็นชิ้นแยกเป็นอิสระกัน เพื่อสะดวกแก่การ นำไปใช้บริโภคเครื่องแช่เยือกแข็ง (freezer) ที่ใช้ เพื่อการแช่เยือกแข็งอาหารทะเล แบบ IQF

- อาหารทะเลแช่แข็งแบบก้อน (block) กุ้งจะถูกนำไปเรียงลงกล่องโลหะ เติมน้ำให้ท่วมก่อนจะถูกนำไปแช่เยือกแข็งด้วยเครื่องแช่เยือกแข็งแบบแผ่น (plate freezer) โดยอาจ

มีอุณหภูมิต่ำถึง -40 องศาเซลเซียส ทั้งนี้ เพื่อให้เนื้อสัตว์ มีอุณหภูมิไม่เกิน -18 องศาเซลเซียส จากนั้นอาหารทะเล จะถูกเคาะออกจากบล็อก อาจมีการเคลือบน้ำแข็ง (ice glazing) เพื่อลดการสูญเสียน้ำก่อนนำไปบรรจุลงถุงและกล่องกระดาษ

(8) การเคลือบ (glazing) ใช้สำหรับผลิตภัณฑ์อาหารทะเลแช่เยือกแข็งแบบก้อน น้ำที่ใช้สำหรับเคลือบหรือใช้เตรียมสารละลายสำหรับเคลือบต้องเป็นน้ำที่สะอาดมีคุณภาพและมาตรฐานตามประกาศกระทรวงสาธารณสุขด้วยน้ำบริโภค

## 2.4 มาตรฐานกึ่งสดและแช่แข็งของประเทศสหรัฐอเมริกา

### 1) การกำหนดเกรด

กระบวนการกำหนดเกรดกึ่งจะมีการประเมินกลิ่น รสชาติ และลักษณะข้อบกพร่องทางกายภาพและความสม่ำเสมอของขนาดของกึ่ง โดยหน่วยตัวอย่างจะสุ่มเก็บจำนวนเป็นไปตามข้อบังคับเกี่ยวกับผลิตภัณฑ์ประมงดังต่อไปนี้

- กึ่งขนาดต่ำกว่า 70 ตัวต่อปอนด์ จะใช้ตัวแทน 2 ปอนด์
- กึ่งขนาด 70-250 ตัวต่อปอนด์ จะใช้ตัวแทน 1 หน่วย
- กึ่งขนาดมากกว่า 250 ตัวต่อปอนด์ จะใช้ตัวแทน 8 ออนซ์

จากนั้นหน่วยตัวอย่างจะได้รับการจัดเกรดตามข้อจำกัด สำหรับข้อบกพร่องโดยสรุปดังตารางที่ 2.2 ดังนี้

ตารางที่ 2.2 หลักเกณฑ์การกำหนดเกรด[6], [7]

การกำหนดเกรด	รสชาติและกลิ่น	จำนวนจุดบกพร่องสูงสุด
เกรด A	ดี	15
เกรด B	ดีพอสมควร	30

### 2) การประเมินผลรสชาติและกลิ่น

- รสชาติและกลิ่นที่ดี (เกรด A) หมายความว่าผลิตภัณฑ์มีรสชาติและมีกลิ่นปกติของกึ่งสด (รสหรือกลิ่นธรรมชาติยอมรับได้) ปราศจากรสชาติและกลิ่นไม่พึงประสงค์ใดๆ

- รสชาติและกลิ่นที่ดีพอสมควร (เกรด B) หมายความว่าผลิตภัณฑ์อาจขาดรสชาติและกลิ่นที่ดีของกึ่งสด แต่ปราศจากรสชาติและกลิ่นไม่พึงประสงค์ใดๆ

## 3) ข้อบกพร่อง

ตัวอย่างการประเมินลักษณะทางกายภาพและข้อบกพร่องให้สอดคล้องกับคำจำกัดความดังต่อไปนี้

## ก) การตรวจสอบในสภาพแช่แข็ง

## (1) ตรวจสอบการคายนํ้าของเนื้อกุ้ง

ต่อคุณภาพ

- การคายนํ้าเล็กน้อย หมายถึง เนื้อกุ้งแทบจะไม่แห้ง ไม่ส่งผลกระทบต่อคุณภาพ

ต่อคุณภาพ

- การคายนํ้าปานกลาง หมายถึง เนื้อกุ้งแห้งอย่างชัดเจน แต่ไม่ส่งผลกระทบต่อคุณภาพ

คุณภาพมาก

- การคายนํ้าที่มากเกินไป หมายถึง เนื้อกุ้งมีการแห้งที่ชัดเจน ส่งผลกระทบต่อ

## ข) การตรวจสอบในสภาพสดหรือละลาย

(1) ขนาดสม่ำเสมอ หมายถึง ระดับของความสม่ำเสมอของตัวอย่างกุ้งที่สังเกตเห็นได้ในขณะตรวจสอบตามมาตรฐาน

(2) จุดต่างดำ หมายถึง การปรากฏตัวของพื้นที่สีดำที่สามารถมองเห็นได้ มีผลต่อความพึงปรารถนาหรือคุณภาพของกุ้ง

## (3) ความเสียหายของชิ้นส่วนกุ้ง

ใช้งานได้

- ความสมบูรณ์ หมายถึง กุ้งต้องมีจำนวนข้อลำตัวครบสมบูรณ์

- การแตกหัก หมายถึง กุ้งต้องไม่ขาดเกินความหนาของลำตัว

- ความเสียหาย หมายถึง กุ้งต้องไม่บดขยี้ หรือเสียหายจนไม่สามารถ

(4) วัสดุที่ไม่ต้องการปะปนกับกุ้ง ได้แก่ ขาเดิน เปลือก หนวด ครีบ และวัสดุที่ไม่เป็นอันตรายที่ไม่ได้เป็นส่วนจากกุ้ง

## (5) หัวกุ้งไม่สามารถยอมรับได้ หมายถึง กุ้งผิดปกติหรือมีอาการของโรค

ที่ต้องการ

## (6) การลอกเปลือกไม่ถูกต้อง หมายถึง การมีหรือไม่มีของเปลือกในรูปแบบ

(7) กุ้งมีการเอาเส้นดำออกไม่ถูกต้อง หมายถึง การปรากฏของเส้นดำหรือลำไส้ที่ต้องกำจัดออกของผลิตภัณฑ์กุ้งที่มีการชักไส้

หนึ่งข้อ

- สำหรับกุ้งขนาดน้อยกว่า 70 ตัวต่อปอนด์ เส้นดำมีความยาวมากกว่า

- สำหรับกุ้งขนาด 71-500 ตัวต่อปอนด์ เส้นด้ามีความยาวมากกว่าสอง  
ซ้อ[6], [7]

## 2.5 เส้นด้ากุ้ง (vein)

เส้นด้ากุ้งหรือทางเดินอาหารตอนปลาย (Protodaeum) เป็นส่วนที่เรียกว่า ลำไส้ เป็นท่อเล็ก ๆ พาดไปทางด้านหลังของลำตัวและไปเปิดออกที่ส่วนท้ายของส่วนท้องเรียกว่า ทวารหนัก[29] ดังแสดงในรูปที่ 2.4 ซึ่งต่อมาจากทางเดินอาหารตอนกลาง (Mesenteron) มีลักษณะเป็น กระเพาะอยู่ถัดจากกระเพาะอาหารไปทางหลัง และฝังอยู่ในต่อมสร้างน้ำย่อย (Hepatopancreas) และมีช่องรับน้ำย่อยด้วย ทางเดินอาหารส่วนนี้จึงทำหน้าที่ในการย่อยอาหาร และทางเดินอาหารตอน หน้า (stomodaeum)ต่อจากทางเดินอาหารตอนกลาง เป็นส่วนของทางเดินอาหารที่ประกอบด้วย ปากอยู่ทางด้านล่างของตอนหัวมีรยางค์รอบปาก 3 คู่ ช่วยในการกินและเคี้ยวอาหารและมีต่อม น้ำลาย (salivary gland) ทำหน้าที่ผลิตน้ำย่อย ต่อจากปากเป็นหลอดอาหารสั้น ๆ นำอาหารเข้าสู่ กระเพาะอาหาร ซึ่งมีลักษณะเป็นถุงใหญ่ ซึ่งอยู่ที่บริเวณส่วนหัวจนถึงอก ทางตอนล่างของกระเพาะมี กล้ามเนื้อหนาช่วยในการบดอาหาร จึงเรียกส่วนนี้ว่า แกสทริกมิลล์ (gastric mill) กระเพาะของกุ้งมี หน้าที่ 2 อย่างคือ เป็นที่พักและบดอาหาร ดังนั้นเส้นด้าซึ่งเป็นส่วนหนึ่งของระบบย่อยอาหารของกุ้ง จะมีสีและความเข้มแตกต่างกันนั้นขึ้นกับพฤติกรรมการกินอาหาร นอกจากกุ้งจะกินอาหารเม็ด สำเร็จรูปแล้ว กุ้งยังสามารถกินอาหารชนิดอื่นที่อยู่ในบ่อได้ เช่น สาหร่าย แพลงก์ตอน สัตว์หน้าดิน เป็นต้น สีที่แตกต่างของเส้นด้าขึ้นอยู่กับอาหารที่กุ้งกินเข้าไป เส้นด้าจะมีสีเข้มขึ้นจากการกินพืชหรือ สัตว์อื่นๆในบ่อ [27] , [30]



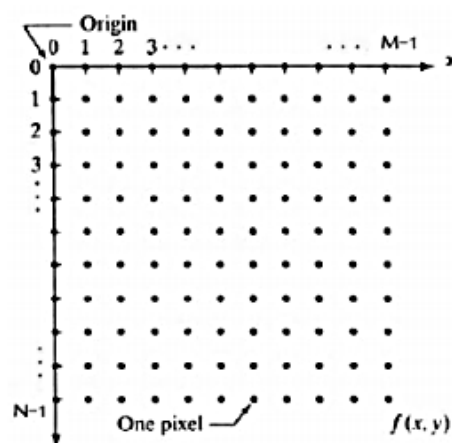
รูปที่ 2.4 เส้นด้าหรือลำไส้ของกุ้ง

## 2.6 ความรู้เบื้องต้นเกี่ยวกับการประมวลผลภาพดิจิทัล

การประมวลผลภาพดิจิทัลเป็นการแปลงข้อมูลภาพให้อยู่ในรูปแบบดิจิทัล (digital format) หรือข้อมูลเชิงตัวเลข โดยวัตถุประสงค์ของการประมวลผลภาพแบ่งออกเป็นสองประเภท คือการปรับปรุงคุณภาพของภาพเพื่อให้มนุษย์สามารถมองเห็นรายละเอียดของภาพมากขึ้น และเพื่อให้คอมพิวเตอร์ตีความหมายของภาพได้ (computer interpretation) [36] ซึ่งจะกล่าวถึงรายละเอียดและทฤษฎีดังต่อไปนี้

### 2.6.1 ภาพดิจิทัล (digital image)

ภาพดิจิทัล คือภาพที่แสดงเป็นอาร์เรย์ (array) หลายมิติ หรือเป็นเมตริกซ์ (matrix) เชิงตัวเลข ภาพดิจิทัลจะแสดงในลักษณะสองมิติ จะมีขนาดความกว้างและความสูงของภาพบนแนวแกน X และแนวแกน Y ส่วนจุดใดๆบนระนาบ XY จะเรียกแทนจุดนั้นว่าพิกเซล (pixel) แต่ละพิกเซลจะแสดงค่าความเข้มแสงของภาพเป็นฟังก์ชัน  $(x,y)$  หรือ  $f(x,y)$  ดังตัวอย่างในรูปที่ 2.5 จากที่กล่าวมาข้างต้นสามารถแทนภาพดิจิทัลด้วยเมตริกซ์ ขนาด  $M \times n$  ดังสมการที่ 2.1



รูปที่ 2.5 ตำแหน่งของพิกเซล[37]

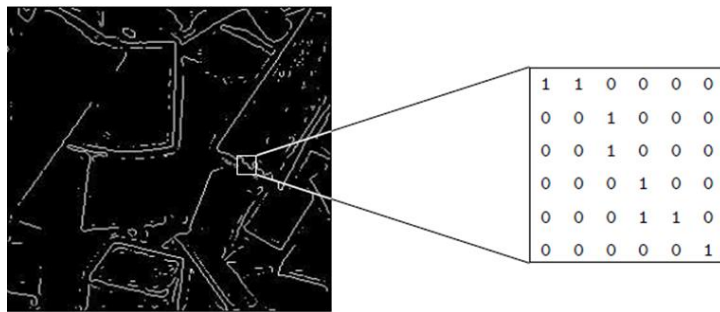
$$f(x, y) = \begin{bmatrix} f(0, 0) & f(0, 1) & \dots & f(0, N-1) \\ f(1, 0) & f(1, 1) & \dots & f(1, N-1) \\ \vdots & \vdots & & \vdots \\ f(M-1, 0) & f(M-1, 1) & \dots & f(M-1, N-1) \end{bmatrix} \quad (2.1)$$



## 2.6.2 ประเภทของภาพ (types of digital images)

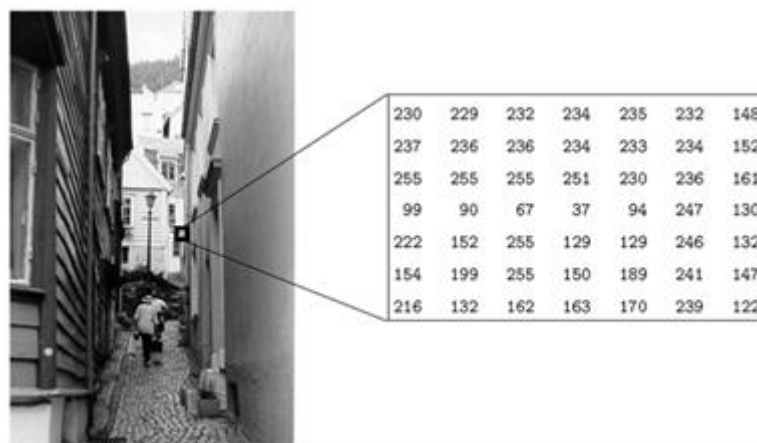
มีการจำแนกภาพออกเป็น 3 ประเภทดังนี้

1) ภาพขาวดำ (binary image, 1-bit) คือภาพที่แต่ละพิกเซลมีค่าหนึ่งบิต สามารถแสดงได้เพียงสองสี คือสีขาวแทนด้วย 1 และสีดำแทนด้วย 0 แสดงดังรูปที่ 2.6



รูปที่ 2.6 ภาพขาวดำ (binary image)[37]

2) ภาพระดับเทา (gray-level image, 8-bit) คือภาพที่มีระดับความเข้มแต่ละพิกเซลสอดคล้องกับความเข้มแสงปกติในระดับสีเทา ซึ่งแต่ละพิกเซลสามารถแสดงเฉดสีจาก 0 (สีดำ) ถึง 255 (สีขาว) [0,255] มีขนาด 8 บิต หรือ 1 ไบต์ แสดงได้ดังรูปที่ 2.7



รูปที่ .2.7 ภาพระดับเทา (gray scale image)[37]

ภาพระดับเทาจะมีการไล่ระดับความเข้มของสีที่อยู่ระหว่างสีขาวและสีดำอย่างต่อเนื่อง โดยที่ค่าแต่ละพิกเซลของภาพจะหมายถึงความเข้มแสงแต่ละตำแหน่งของพิกเซลที่อยู่ในรูป

ระดับสีเทา การแปลงภาพจากระบบสีอาร์จีบี (RGB) เป็น ระดับสีเทา (Gray Scale) จะใช้สมการที่ 2.2 ในการแปลง[38]

$$Gr(x,y) = 0.2989R(x,y)+0.5870G(x,y)+0.1140B(x,y) \quad (2.2)$$

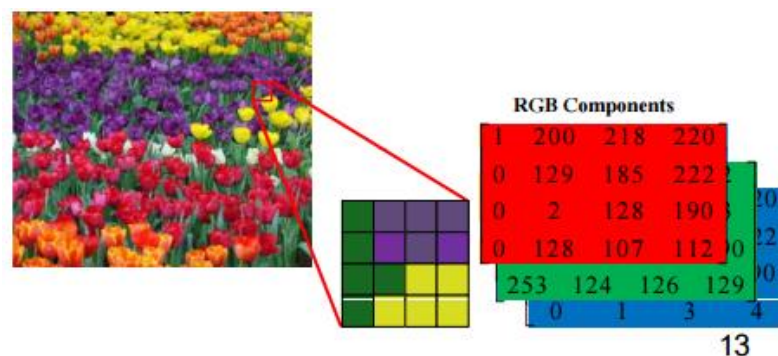
โดย  $Gr(x,y)$  แทน ค่าความเข้มของระดับสีเทา ณ ตำแหน่ง  $(x,y)$

$R(x,y)$  แทน ค่าความเข้มของสีแดง ณ ตำแหน่ง  $(x,y)$

$G(x,y)$  แทน ค่าความเข้มของสีเขียว ณ ตำแหน่ง  $(x,y)$

$B(x,y)$  แทน ค่าความเข้มของสีน้ำเงิน ณ ตำแหน่ง  $(x,y)$

3) ภาพสี (color image) หรือ 24-bit RGB color image คือภาพสีที่แต่ละพิกเซลมีส่วนประกอบของเวกเตอร์แทนสีแดง (red), สีเขียว (green) และสีน้ำเงิน (blue) โดยแต่ละสีสามารถแสดงความเข้มแสงได้ในช่วง  $[0,255]$  ทำให้แต่ละพิกเซลสามารถแสดงสีที่เป็นไปได้ทั้งหมด  $256^3$  เท่ากับ 16,777,216 แต่ละพิกเซลมีขนาด 24 บิต แสดงได้ดังรูปที่ 2.8

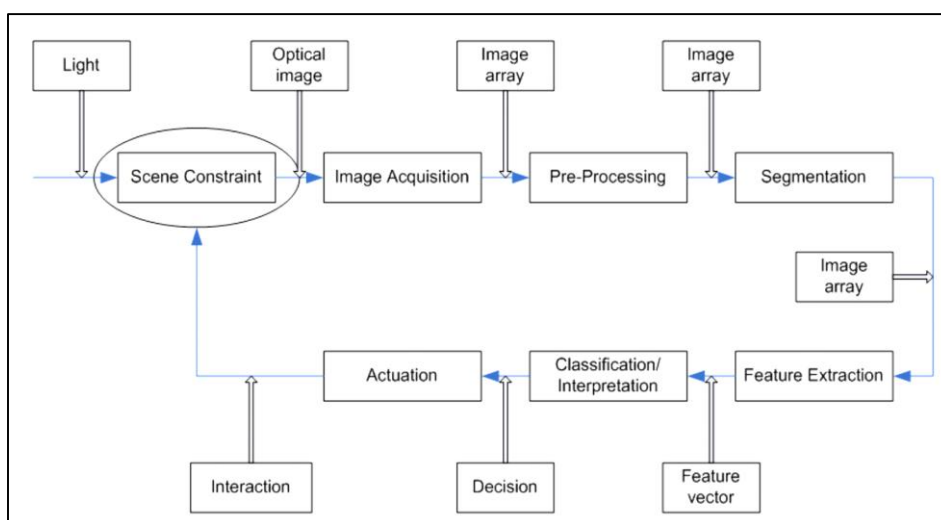


รูปที่ 2.8 ภาพสี (color image)[37]

## 2.7 การประมวลผลภาพ (image processing)

การประมวลผลภาพ (image processing) หมายถึง การนำภาพมาประมวลผลหรือคิดคำนวณด้วยคอมพิวเตอร์ เพื่อให้ได้ข้อมูลที่เราต้องการทั้งในเชิงคุณภาพและปริมาณโดยมีขั้นตอนต่างๆที่สำคัญ คือ การทำให้ภาพมีความคมชัดมากขึ้น การกำจัดสัญญาณรบกวนออกจากภาพ การแบ่งส่วนของวัตถุที่เราสนใจออกมาจากภาพ เพื่อนำภาพวัตถุที่ได้ไปวิเคราะห์หาข้อมูลเชิงปริมาณ เช่น ขนาด รูปร่าง และทิศทางการเคลื่อนของวัตถุในภาพ จากนั้นเราสามารถนำข้อมูลเชิงปริมาณเหล่านี้ไปวิเคราะห์ และสร้างเป็นระบบ เพื่อใช้ประโยชน์ในงานด้านต่างๆ เช่น ระบบตรวจสอบคุณภาพของผลิตภัณฑ์ในกระบวนการผลิตของโรงงานอุตสาหกรรม ระบบคัดแยกเกรด

ระบบตรวจสอบชิ้นงานด้วยภาพแบบอัตโนมัติ (Automated visual inspection system) เป็นการนำเอาประมวลผลด้วยภาพไปประยุกต์ใช้ในงานอุตสาหกรรมเพื่อใช้ตรวจสอบคุณภาพของผลิตภัณฑ์ ซึ่งผลของการตรวจสอบดังกล่าวอาจใช้เพื่อคัดแยกงานดีออกจากงานเสีย หรือใช้เพื่อคัดเลือกเกรดของชิ้นงาน และจากความต้องการจากด้านของอุตสาหกรรมที่ต้องการระบบที่สามารถทำงานได้รวดเร็วมากที่สุด ทำให้การทำงานของระบบอัตโนมัติ จะต้องใช้เวลาที่สั้นที่สุด ดังนั้นวิธีการคำนวณต่างๆจะต้องใช้เวลาให้น้อยที่สุด นอกจากนั้นความรู้ต่างๆที่มีอยู่ในผู้ที่ปฏิบัติงานกับผลิตภัณฑ์นั้นๆมาก่อน จะต้องถูกนำมาใช้เพื่อให้ระบบสามารถตัดการคำนวณที่ไม่จำเป็นออกไปให้ได้มากที่สุด และสามารถอธิบายขั้นตอนพื้นฐานของการประมวลผลภาพด้วยคอมพิวเตอร์[39]ในรูปที่ 2.9 ได้ดังนี้



รูปที่ 2.9 ขั้นตอนพื้นฐานของการประมวลผลภาพด้วยคอมพิวเตอร์[39]

### 2.7.1 การจัดการสภาพแวดล้อม (scene constraint)

จุดมุ่งหมายหลักของการจัดการสภาพแวดล้อมของระบบตรวจสอบชิ้นงานด้วยภาพแบบอัตโนมัติ คือ เพื่อลดความซับซ้อนในการประมวลผลให้มากที่สุด เนื่องจากการที่ความสามารถการมองเห็นและรับรู้ของอุปกรณ์ประมวลผลที่มีอยู่ในปัจจุบันมีอยู่อย่างจำกัด และไม่เทียบเท่ากับความสามารถของมนุษย์ ดังนั้นจึงต้องช่วยลดความยุ่งยากของการประมวลผล ทั้งนี้เพื่อให้อุปกรณ์ประมวลผลใช้เวลาส่วนใหญ่ไปกับงานที่ไม่ซับซ้อนและเท่าที่มีความจำเป็นเท่านั้น ได้แก่

1) การจัดการกับชิ้นงาน ซึ่งหากไม่จัดการกับการวางตัวของชิ้นงาน อุปกรณ์ประมวลผลต้องหาทิศทางของชิ้นงานเอง

2) ระยะระหว่างกล้องหรือเลนส์ถึงชิ้นงาน ตัวแปรเหล่านี้จะเป็นตัวกำหนดขนาดของชิ้นงานที่ระบบมองเห็น

3) การจัดการเรื่องแสง แสงจัดเป็นองค์ประกอบสำคัญมาก เนื่องจากการมองเห็นภาพของระบบอัตโนมัติ นั้น เกิดจากการที่มีแสงมาตกกระทบวัตถุ แล้วสะท้อนผ่านเลนส์มาเข้าตัวเซนเซอร์รับภาพของกล้องที่ใช้กับระบบอัตโนมัติ ซึ่งการจัดการเกี่ยวกับแสงนั้น จำเป็นจะต้องพิจารณาทั้งเรื่องการเลือกใช้แหล่งกำเนิดแสง การกระเจิงของแสง และคุณสมบัติอื่นๆ สำหรับการตรวจสอบชิ้นงานโดยทั่วไปแล้วจะทำการติดตั้งแหล่งกำเนิดแสงไว้ที่ด้านเดียวกับตัวกล้องแล้วส่องไปที่วัตถุที่ต้องการจับภาพ ซึ่งเรียกกันว่า front lighting อย่างไรก็ตามจะพบว่า สำหรับงานตรวจสอบชิ้นงานบางอย่าง อาจจำเป็นต้องใช้การส่องแสงมาจากด้านหลังของวัตถุที่ใช้พิจารณาที่เรียกกันว่า back lighting แล้วใช้ภาพที่ได้ทั้ง 2 ภาพ ดังรูปที่ 2.10



ภาพที่มีแหล่งกำเนิดแสงจากด้านหน้า  
(front lighting)

ภาพที่มีแหล่งกำเนิดแสงจากด้านหลัง  
(back lighting)

รูปที่ 2.10 ภาพกุ้ง

นอกจากการจัดการสภาพแวดล้อมในการทำงานให้กับระบบตรวจสอบชิ้นส่วนจากภาพแบบอัตโนมัติแล้ว งานบางประเภทอาจจะต้องมีการใช้ภาพจากกล้องหลายตัว เพื่อใช้ตรวจสอบชิ้นงานจากหลายมุมมอง บางกรณีอาจจะเป็นการใช้กล้องเพียงตัวเดียว แต่ตัวกล้องสามารถเคลื่อนที่ไปตามส่วนต่างๆของชิ้นงานได้ และสำหรับบางกรณีอาจจะมีการเก็บภาพของวัตถุเดียวกัน ที่ได้จากแหล่งกำเนิดแสงหลายแหล่ง

### 2.7.2 การดึงข้อมูลภาพ (image acquisition)

กระบวนการดึงข้อมูลภาพคือ กระบวนการที่เริ่มตั้งแต่การถ่ายภาพโดยกล้อง จนถึงการดึงภาพซึ่งเป็นข้อมูลที่อยู่ในกล้องมาสู่คอมพิวเตอร์หรืออุปกรณ์ประมวลผล เพื่อที่จะได้ประมวลผลและตัดสินใจสั่งงานจากผลที่ได้ต่อไป กระบวนการดังกล่าวมีรายละเอียดปลีกย่อยที่สำคัญ ดังนี้

ประเภทของกล้องที่ใช้ในงานตรวจสอบชิ้นส่วนในเชิงอุตสาหกรรม กล้องที่ใช้กับงานตรวจสอบชิ้นส่วนในปัจจุบันนั้น จะเป็นกล้องดิจิทัลซึ่งใช้อุปกรณ์สารกึ่งตัวนำที่เรียกกันว่า เซ็นเซอร์รับภาพ (Image sensor) เพื่อใช้ในการรับภาพ เซ็นเซอร์ดังกล่าวมีขนาดเล็กมากเท่าเล็บมือคนเท่านั้น ซึ่งจะประกอบด้วยไดโอดที่มีความไวต่อแสงเรียงตัวกันอยู่เป็นจำนวนมาก และในทันทีทันใดที่แสงมีการตกกระทบไดโอดเหล่านี้ ไดโอดแต่ละตัวจะทำการจดจำความเข้มแสงหรือความสว่างของแสงที่ตกกระทบไดโอดแต่ละตัวไว้ โดยปริมาณประจุไฟฟ้าที่สะสมอยู่ในตัวไดโอดซึ่งแปรผันกับแรงดันตกคร่อมตัวไดโอดนั้น จะเพิ่มขึ้นตามความเข้มของแสงที่ตกกระทบ ซึ่งความเข้มแสงที่ได้จดจำไว้ในไดโอดแต่ละตัว จะถูกแปลงให้อยู่ในรูปข้อมูลที่เป็นดิจิทัลและเก็บไว้ในหน่วยความจำที่อยู่ในตัวกล้อง เพื่อรอส่งต่อไปให้อุปกรณ์ที่อยู่ภายนอกกล้องต่อไป

สำหรับไดโอดนี้เรียกกันว่า เซลล์รับภาพ (pixel) ซึ่งหนึ่งเซลล์รับภาพจะให้ค่าความเข้มแสงที่ตกกระทบเพียงค่าหนึ่งเท่านั้น โดยทั่วไปค่าที่ได้จากเซลล์รับภาพจะมีค่าระหว่าง 0-255 เท่านั้น (ช่วงข้อมูลดังกล่าว สามารถแทนด้วยข้อมูลขนาด 1 byte หรือ 8 บิต ที่จะให้ความละเอียด 28 หรือ 256 ระดับ ซึ่งเป็นความละเอียดของกล้องที่สามารถพบเห็นได้ทั่วไปในท้องตลาด อย่างไรก็ตาม จะมีกล้องบางประเภทที่ให้ค่าความเข้มแสงที่มีความละเอียดสูงถึง 16 บิตเลยทีเดียว) โดยหากค่าที่ได้มีค่าเท่ากับ 0 แสดงว่าที่เซลล์รับภาพนั้น มีความเข้มแสงต่ำสุดหรือเป็นด้านมืด และหากมีค่าเท่ากับ 255 ก็แสดงว่าที่เซลล์รับภาพที่ตำแหน่งนั้นมีความเข้มแสงสูงสุดหรือเป็นด้านสว่างดังแสดงไว้ในรูปที่ 2.11



รูปที่ 2.11 ความเข้มแสงเทียบกับค่าที่ได้จากเซลล์รับภาพ[39]

เมื่อพิจารณากระบวนการทั้งหมด จะพบว่าเซลล์รับภาพแต่ละเซลล์นั้นจะให้ค่าความเข้มแสงออกมาเป็นตัวเลขที่อยู่ระหว่าง 0 ถึง 255 เท่านั้น ซึ่งทำให้ได้ภาพที่เป็นโทนขาวดำหรือ

เรียกว่า gray scale image เท่านั้น และกล้องสีใช้เซลล์รับภาพเช่นกัน โดยทำการแยกสีหลัก 3 สี ได้แก่ สีแดง เขียว และน้ำเงินออกจากกัน โดยการติดตั้งตัวกรองแสงสี (filter) แต่ละสีไว้หน้าเซลล์รับภาพแล้วทำการวัดความเข้มของแต่ละสีนั่นเอง

ภาพที่อุปกรณ์ประมวลผลมองเห็น หลักการทำงานของกล้องก็เป็นเช่นเดียวกับระบบการมองเห็นของมนุษย์ นั่นคือ ภาพเกิดจากการที่มีแสงตกกระทบวัตถุแล้วมีแสงสะท้อนจากวัตถุ ผ่านเลนส์เข้ามาตกกระทบเซ็นเซอร์รับภาพ (image sensor) ของกล้อง ซึ่งประกอบด้วยเซลล์รับภาพ (pixel) จำนวนมาก เซลล์รับภาพแต่ละเซลล์จะทำหน้าที่แปลงความเข้มแสงสำหรับกรณีที่เป็นกล้องขาวดำหรือแปลงความเข้มสีของแสงสีแดง เขียว และน้ำเงินสำหรับกรณีของกล้องที่ใช้ถ่ายภาพสี ให้อยู่ในรูปของค่าสัญญาณแรงดันไฟฟ้า ซึ่งจะถูกละเปลี่ยนไปเป็นสัญญาณดิจิทัลด้วยตัวแปลงสัญญาณอนาล็อกเป็นดิจิทัลอีกทีหนึ่ง อย่างไรก็ตามการทำงานของเซลล์รับภาพของกล้องจะแตกต่างจากเซลล์รับภาพของมนุษย์อยู่ 2 ประการด้วยกัน คือ

- จำนวนเซลล์รับภาพที่ประกอบกันขึ้นมาเป็นเซ็นเซอร์รับภาพของกล้องนั้น มีจำนวนน้อยกว่าของมนุษย์เป็นอย่างมาก ทำให้ภาพที่ได้จากกล้องนั้นมีความละเอียดน้อยกว่าของมนุษย์เป็นอย่างมาก ภาพที่ได้จากกล้องจึงเกิดการสุ่มจับภาพจริงด้วยจำนวนที่จำกัดของเซลล์รับภาพ (Spatial sampling) นั่นเอง

- ค่าความเข้มแสงที่ได้จากเซลล์รับภาพ (หรือความเข้มสีในกรณีของกล้องถ่ายภาพสี) ของเซ็นเซอร์รับภาพที่อยู่ในกล้องนั้น จะเป็นค่าไม่ต่อเนื่อง (discrete value) เนื่องจากการทำงานของอุปกรณ์ดิจิทัลซึ่งจะเป็นการสุ่มขนาดของความเข้มแสงที่ตกกระทบ (Amplitude sampling) ไม่เหมือนกับของมนุษย์ที่มีความต่อเนื่อง เนื่องจากการทำงานของสารเคมีที่อยู่ในเซลล์รับภาพ

### 2.7.3 การประมวลผลภาพเบื้องต้น (pre-processing)

การประมวลผลภาพมีด้วยกันหลากหลายกระบวนการด้วยกัน เช่น การลดทอนสัญญาณรบกวนที่ปรากฏขึ้นในภาพดังรูปที่ 2.12 การตรวจจับขอบของวัตถุที่อยู่ในภาพดังรูปที่ 1.13 การแปลงคุณสมบัติทางกายภาพของภาพ การย่อและขยายภาพ การแปลงสีดังรูปที่ 1.14 การวิเคราะห์ภาพในเชิงความถี่ และการบีบอัดข้อมูลภาพ เป็นต้น



ก) ภาพต้นฉบับ



ข) ภาพที่กำจัดสัญญาณรบกวน

รูปที่ 2.12 การกำจัดสัญญาณรบกวน[40]



ก) ภาพต้นฉบับ



ข) ภาพที่มีการจับขอบวัตถุ

รูปที่ 2.13 การหาขอบภาพ[40]



ก) ภาพต้นฉบับ



ข) ภาพที่แปลงภาพเป็นภาพระดับเทา

รูปที่ 2.14 การแปลงภาพสีเป็นภาพระดับเทา[40]

ระเบียบวิธี (algorithm) ของวิธีประมวลผลภาพบางอย่างก็ไม่เหมาะสมที่จะนำมาใช้กับงานตรวจสอบชิ้นส่วนแบบอัตโนมัติ เนื่องจากการนำไปใช้ในทางอุตสาหกรรมที่ต้องการการทำงานที่รวดเร็วที่สุด ดังนั้นจึงมีความจำเป็นอย่างยิ่งที่จะต้องเลือกใช้ระเบียบวิธีเฉพาะที่



ง่ายและใช้เวลาในการทำงานน้อยที่สุด ซึ่งเงื่อนไขเหล่านี้จะสามารถเป็นจริงได้ ก็ด้วยการจัดสภาพแวดล้อมในการจับภาพที่ดี นอกจากนั้นแล้ว ความรู้เกี่ยวกับผลิตภัณฑ์นั้นซึ่งมีอยู่แล้วในผู้ปฏิบัติงาน ก็ควรนำมาใช้เพื่อช่วยให้ระบบสามารถทำงานให้ได้เร็วที่สุดด้วยเช่นกัน ยกตัวอย่างเช่น

การระบุบริเวณที่ต้องการตรวจสอบ (Region of Interest, ROI) โดยผู้ปฏิบัติงาน รวมทั้งประเภทการตรวจสอบที่ใช้กับบริเวณนั้นๆ จะทำให้โปรแกรมสามารถตัดการคำนวณของบริเวณที่ไม่เกี่ยวข้องออกไป ทำให้โปรแกรมสามารถทำงานได้เร็วขึ้น

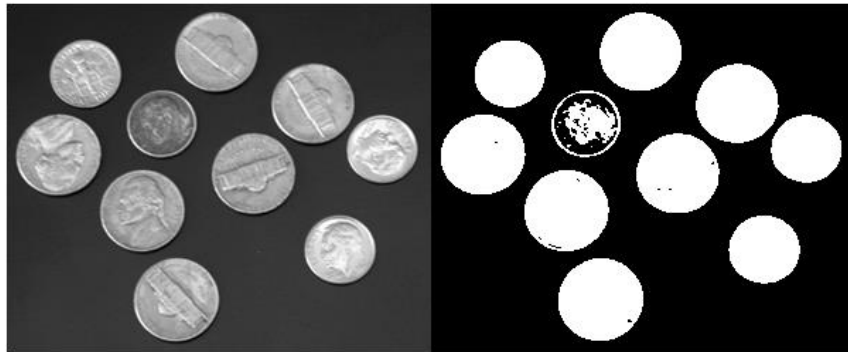
การนำความรู้เรื่องรูปร่างของผลิตภัณฑ์มาใช้ เช่น การตรวจสอบบริเวณขอบของผลิตภัณฑ์ที่เป็นส่วนของเส้นตรง บริเวณที่เป็นวงกลม จะพบว่า การตรวจสอบประเภทนี้จะต้องทำการตรวจจับส่วนของเส้นตรงหรือการตรวจจับวงกลมก่อนที่จะทำการพิจารณาคุณสมบัติอื่นๆ เป็นต้น

#### 2.7.4 การแยกบริเวณ (segmentation)

การแบ่งย่อยส่วนประกอบภาพ ใช้ในการแยกองค์ประกอบต่างๆของรูปภาพออกจากกันตามลักษณะสำคัญที่พิจารณา ขั้นตอนนี้เป็นขั้นตอนที่สำคัญของการวิเคราะห์ภาพ ดังนั้นประโยชน์ของการแบ่งย่อยส่วนประกอบภาพคือ ลดจำนวนข้อมูลในรูปภาพที่ไม่จำเป็นออกจากการวิเคราะห์ เนื่องจากการแบ่งย่อยส่วนประกอบของภาพเป็นการแบ่งพื้นที่ของส่วนที่สนใจ และเป็นการจัดระเบียบข้อมูลในภาพให้เป็นกลุ่มได้ดีขึ้น ซึ่งง่ายต่อการนำภาพไปวิเคราะห์ เราสามารถแบ่งประเภทการแบ่งย่อยส่วนประกอบภาพได้ 2 แบบคือ

1) การแยกบริเวณโดยการใช้ค่าเทรชโฮลด์ (threshold) ซึ่งเป็นค่าที่เป็นจำนวนเต็มที่มีค่าระหว่าง 0 ถึง 255 เช่นเดียวกับค่าความเข้มแสงของพิกเซลที่อยู่ในภาพสีเทา สำหรับการแยกบริเวณโดยการใช้ค่าเทรชโฮลด์ นั้นจะเป็นการแปลงภาพสีเทาให้เปลี่ยนเป็นภาพที่มีเพียงสองระดับ โดยมีเงื่อนไขว่าถ้าค่าความเข้มแสงที่พิกเซลตำแหน่งใดมีค่าต่ำกว่าหรือเท่ากับเทรชโฮลด์ ให้ค่าพิกเซลในตำแหน่งนั้นมีค่าเป็น 0 หรือเปลี่ยนเป็นด้านมืดไปและถ้าค่าความเข้มแสงที่พิกเซลตำแหน่งใดมีค่าสูงกว่าค่าเทรชโฮลด์ ให้ค่าพิกเซลในตำแหน่งนั้นมีค่าเป็น 255 หรือเปลี่ยนเป็นด้านสว่างไป ดังรูปที่ 2.15





รูปที่ 2.15 การแยกบริเวณโดยการใช้ค่าเทรชโฮลด์ (threshold)[40]

2) การแยกบริเวณโดยการใช้ขอบของวัตถุ (edge base segmentation) สำหรับการแบ่งบริเวณวิธีนี้ ต้องคำนวณหาขอบของวัตถุเสียก่อน ซึ่งขอบในความหมายของการประมวลผลภาพแบบดิจิทัลคือ พิกเซลที่มีการเปลี่ยนแปลงความเข้มแสงเกินค่าที่กำหนด ซึ่งสามารถตรวจจับได้โดยการใช้ตัวตรวจจับขอบ (edge detector) ซึ่งคำนวณจากการเปลี่ยนแปลงของความเข้มในตำแหน่งจุดภาพที่ใกล้เคียงกัน ดังรูปที่ 2.16



รูปที่ 2.16 การแยกบริเวณโดยการใช้ขอบของวัตถุ (edge base segmentation)[40]

### 2.7.5 การคำนวณหาคุณสมบัติของวัตถุ (feature extraction)

การแยกลักษณะเฉพาะของภาพเป็นการแยกหรือสกัดเอาข้อมูลที่สำคัญของภาพออกมา ซึ่งลักษณะเฉพาะของภาพเป็นคุณสมบัติที่สามารถหาได้โดยใช้ขั้นตอนวิธีการประมวลผลภาพ (image processing) โดยที่ลักษณะเฉพาะพื้นฐานของภาพประกอบด้วย 3 ส่วนคือ สี รูปร่าง และพื้นผิว

1) สี (color) เป็นลักษณะเฉพาะของภาพที่มีบทบาทสำคัญในระบบค้นคืนภาพ เช่น ฮิสโตแกรมสีซึ่งเป็นลักษณะเฉพาะของสีที่ถูกนำมาใช้บ่อยๆ เนื่องจากสีเป็นสิ่งที่สามารถมองเห็น

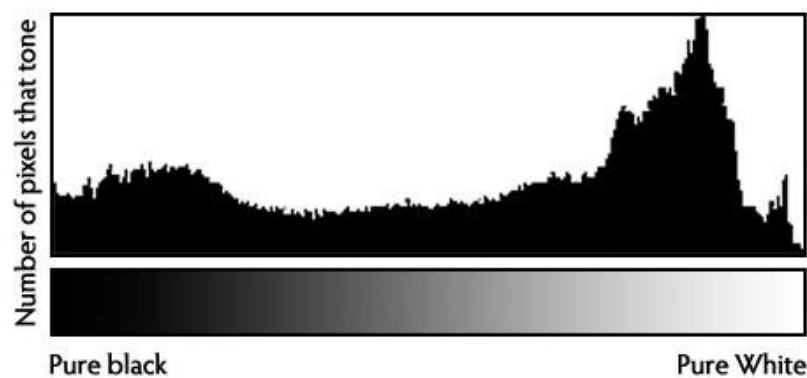
ได้ง่าย และเป็นสิ่งแรกที่สามารถสังเกตเห็นได้จากการมองภาพนอกจากนี้สียังสามารถใช้ในการแยกแยะกลุ่มของภาพออกตามเนื้อหาได้เป็นอย่างดี เช่น สีฟ้าของน้ำทะเล สีแดงของดอกไม้ สีเขียวของต้นไม้ เป็นต้น

2) รูปร่าง (shape) เป็นลักษณะเฉพาะของภาพที่ใช้อธิบายถึงรูปร่างและลักษณะรวมถึงขนาดของวัตถุภายในภาพซึ่งทำให้สามารถแยกวัตถุออกจากพื้นหลังหรือแยกแยะระหว่างวัตถุที่มีรูปร่างแตกต่างกันออกจากกันได้

3) พื้นผิว (texture) เป็นลักษณะเฉพาะที่ใช้อธิบายความหยาบความละเอียดหรือความซับซ้อนของวัตถุภายในภาพซึ่งแต่ละภาพอาจจะประกอบด้วยวัตถุที่มีลักษณะพื้นผิวที่แตกต่างกันออกไปการวิเคราะห์พื้นผิวจะช่วยให้สามารถแยกแยะความแตกต่างของวัตถุได้ดียิ่งขึ้นการค้นคืนภาพที่ใช้พื้นผิวเป็นลักษณะเฉพาะของภาพส่วนใหญ่จะถูกนำไปใช้ในการค้นหาภาพจากกลุ่มภาพพื้นผิว เช่น ชุดภาพพื้นผิวของหินชุดภาพพื้นผิวของใบไม้ เป็นต้น

นอกจากลักษณะเฉพาะพื้นฐานของภาพทั้ง 3 ส่วนที่กล่าวมาแล้วนั้น ในงานวิจัยนี้ได้อาศัยลักษณะของฮิสโตแกรมโดยใช้ข้อมูลพื้นฐาน เช่น ค่าเฉลี่ย ค่าต่ำสุด ค่าสูงสุด ส่วนเบี่ยงเบนมาตรฐาน เป็นต้น ซึ่งมีรายละเอียดดังต่อไปนี้

ฮิสโตแกรม (Histogram) เป็นกราฟแท่งที่แสดงถึงความถี่หรือจำนวนของพิกเซลที่กระจายอยู่ ณ ช่วงค่าความสว่างต่างๆของภาพ โดยไม่สนใจว่าตำแหน่งของพิกเซลนั้นจะอยู่ส่วนใดของภาพ แกน X เป็นระดับความสว่างที่แบ่งระดับเป็น 256 ระดับ มีค่าตั้งแต่ 0-255 ถ้าระดับสีเทามีค่าต่ำ หมายถึงมีความสว่างน้อยจะมองเห็นเป็นสีดำ และค่าระดับสีเทามาก หมายถึงมีความสว่างมากจะมองเห็นเป็นสีขาว แกน Y ของกราฟแสดงจำนวนพิกเซลในแต่ละความระดับสีเทา ซึ่งเป็นค่าสัมพัทธ์[41] ดังรูปที่ 2.17



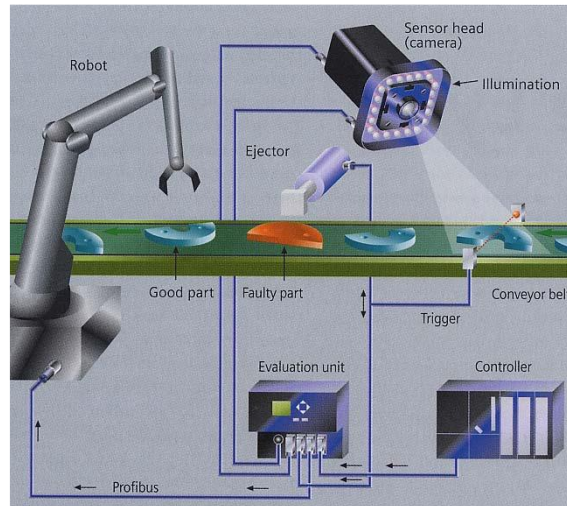
รูปที่ 2.17 ฮิสโตแกรมภาพระดับเทา[42]

### 2.7.6 การจำแนกวัตถุและแปลความหมาย (classification/interpretation)

การจำแนก (classification) คือ กระบวนการจัดกลุ่มให้วัตถุที่กำลังพิจารณาอยู่นั้นว่าเป็นวัตถุที่อยู่ในกลุ่มที่ 1, 2 หรือกลุ่มอื่นๆ โดยอาศัยข้อมูลที่ได้จากการวัดหรือการคำนวณซึ่งเป็นคุณสมบัติของวัตถุเปรียบเทียบกับคุณสมบัติของวัตถุตัวอย่างที่อยู่ในกลุ่มแต่ละกลุ่ม ซึ่งผลการทำงานจะทำให้สามารถทราบว่าเป็นวัตถุชนิดใด ซึ่งจะทำให้สามารถตีความหมาย (interpretation) และตัดสินใจสั่งการส่วนเคลื่อนไหวต่างๆได้

### 2.7.7 กลไกเคลื่อนไหว (actuation)

กระบวนการนี้โปรแกรมจะสั่งการส่วนกลไกเคลื่อนไหวต่างๆ ให้กระทำการบางอย่างกับผลิตภัณฑ์ที่ผ่านกระบวนการก่อนหน้า ซึ่งในส่วนนี้งานหลักๆจะเป็นการติดต่อและสั่งงานระหว่างอุปกรณ์ประมวลผลผลและ programmable logic control ที่สามารถใช้สั่งการส่วนเคลื่อนไหวต่างๆ เช่น มอเตอร์ แขนหุ่นยนต์ หรืออื่นๆ ดังรูปที่ 2.18



รูปที่ 2.18 การทำงานจริงที่ประกอบด้วยกลไกต่างๆ[43]

## 2.8 สถิติที่ใช้ในการเตรียมข้อมูล

### 2.8.1 การทดสอบค่าเฉลี่ยกรณีกลุ่มตัวอย่าง 2 กลุ่ม

ในการทดสอบค่าเฉลี่ยกรณีกลุ่มตัวอย่าง 2 กลุ่มนั้นจะพิจารณาว่ากลุ่มตัวอย่างทั้งสองกลุ่มเป็นอิสระจากการหรือไม่ เพื่อเลือกใช้สูตรของสถิติทดสอบให้ถูกต้อง นอกจากนี้ยังพิจารณาอีกว่าความแปรปรวนของประชากรของกลุ่มตัวอย่างเท่ากันหรือไม่ ซึ่งในการใช้สถิติ t-test ทดสอบกรณีกลุ่มตัวอย่าง 2 กลุ่ม ที่เป็นอิสระต่อกันนั้นมีสูตรที่ใช้ทดสอบอยู่ 2 สูตรด้วยกัน กล่าวคือ สูตรที่ใช้ในกรณีความแปรปรวนของประชากร 2 กลุ่มมีค่าเท่ากัน ( $\sigma_1^2 = \sigma_2^2$ ) หรือในกรณีกลุ่มตัวอย่างแต่ละกลุ่มมีจำนวนเท่ากัน (t-test แบบ Pooled variance) และสูตรที่ใช้ในกรณีความแปรปรวนของประชากร 2 กลุ่ม มีค่าไม่เท่ากัน ( $\sigma_1^2 \neq \sigma_2^2$ ) (t-test แบบ Separated variance) ดังนั้นเมื่อผู้วิจัยจะใช้ t-test กรณีดังกล่าวจะต้องทำการทดสอบก่อนว่าความแปรปรวนของประชากรแต่ละกลุ่มมีค่าเท่ากันหรือไม่โดยใช้ F-test เพื่อจะได้เลือกใช้สูตรของ t-test ได้อย่างถูกต้องและเหมาะสมต่อไป[45]

1) กรณีกลุ่มตัวอย่างเป็นอิสระต่อกัน

ก) เมื่อสุ่มตัวอย่างขนาด  $n_1$  และ  $n_2$  มาโดยอิสระจากกัน มีการแจกแจงแบบปกติ ที่มีค่าเฉลี่ย  $\mu_1$  และ  $\mu_2$  ความแปรปรวน  $\sigma_1^2$  และ  $\sigma_2^2$  ซึ่งไม่ทราบค่า แต่ทราบว่า  $\sigma_1^2 = \sigma_2^2$  โดย  $n_1$  และ  $n_2$  น้อยกว่า 30 ใช้สูตร t - test (t-test แบบ Pooled variance)

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S_p^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}} \quad df = n_1 + n_2 \quad (2.6)$$

เมื่อ

$$S_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2} \quad (2.7)$$

ข) เมื่อสุ่มตัวอย่างขนาด  $n_1$  และ  $n_2$  มาโดยอิสระจากกัน มีการแจกแจงแบบปกติ ที่มีค่าเฉลี่ยเท่ากับ  $\mu_1$  และ  $\mu_2$  ความแปรปรวนเท่ากับ  $\sigma_1^2$  และ  $\sigma_2^2$  ซึ่งไม่ทราบค่าแต่ทราบที่  $\sigma_1^2 \neq \sigma_2^2$  ใช้สูตร t - test (t-test แบบ Separated variance)

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad ; \quad df = \frac{\frac{s_1^2 + s_2^2}{n_1 + n_2}}{\frac{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}{n_1 + n_2 - 2}} \quad (2.8)$$

เนื่องจากการทดสอบทั้ง 2 กรณีข้างต้นเกี่ยวข้องกับการทราบค่าของความแปรปรวน  $\sigma_1^2$  และ  $\sigma_2^2$  ว่า เท่ากันหรือไม่ ดังนั้นในการวิเคราะห์ข้อมูลกรณีตัวอย่าง 2 กลุ่ม เราจึงจำเป็นต้องทำการทดสอบความเท่ากันของความแปรปรวนโดยใช้สถิติทดสอบ F-test ก่อนเพื่อเลือกใช้ให้ถูกต้อง ดังนี้

$$F = \frac{S_{\max}^2}{S_{\min}^2} ; \quad (2.9)$$

$df_1 = n_1 - 1$  เมื่อ  $n_1 =$  จำนวนกลุ่มตัวอย่างที่มีค่า  $S^2$  มีค่ามาก

$df_2 = n_2 - 1$  เมื่อ  $n_2 =$  จำนวนกลุ่มตัวอย่างที่มีค่า  $S^2$  มีค่าน้อย

## 2) กรณีกลุ่มตัวอย่างมีความสัมพันธ์กัน

การเปรียบเทียบความแตกต่างของคะแนนเฉลี่ย 2 กลุ่มที่มีความสัมพันธ์กัน (ไม่อิสระจากกัน) เช่น ในกรณีที่ทำการศึกษาที่มีลักษณะเป็นคู่กัน เช่น ฝาแฝด หรือคนกลุ่มเดียวแต่มีการทดสอบสองครั้ง เช่น ทดสอบก่อนการทดลองและทดสอบหลังการทดลอง โดยใช้ค่าแจกแจง t - test แบบ Dependent Samples ดังนี้

$$t = \frac{\sum D}{\sqrt{\frac{N \sum D^2 - (\sum D)^2}{N - 1}}} ; \quad df = n - 1 \quad (2.10)$$

เมื่อ	$t$	แทน	ค่าสถิติที่ใช้ในการพิจารณาใน t - distribution
	$D$	แทน	ความแตกต่างของคะแนนแต่ละคู่
	$N$	แทน	จำนวนคู่ของคะแนนหรือจำนวนนักเรียน
	$\sum D$	แทน	ผลรวมทั้งหมดของผลต่างคะแนนก่อนและหลังการทดลอง

$\sum D^2$  แทน ผลรวมของกำลังสองของผลต่างคะแนนก่อนและหลังทดลอง

## 2.8.2 การวิเคราะห์ปัจจัย

การวิเคราะห์ปัจจัยเป็นวิธีการทางสถิติที่วิเคราะห์ตัวแปรหลายตัววิธีหนึ่งที่นิยมใช้ในการลดจำนวนตัวแปรให้น้อยลงเพื่อให้ง่ายต่อการเข้าใจ โดยศึกษาโครงสร้างความสัมพันธ์ระหว่างตัวแปรและสร้างตัวแปรใหม่ที่เรียกว่า ปัจจัย (factor) โดยปัจจัยที่สร้างขึ้นจะประกอบด้วยรายละเอียดหรือความแปรผันของตัวแปรเดิมหลายๆตัว หรือเป็นการนำตัวแปรที่มีความสัมพันธ์กันมากหรือมีความร่วมกันมากมารวมกันไว้เป็นปัจจัยเดียวกัน ส่วนตัวแปรที่มีความสัมพันธ์น้อยหรือมีความร่วมกันน้อยหรือไม่มีความสัมพันธ์กันไว้คนละปัจจัย

การวิเคราะห์ปัจจัยเป็นวิธีที่ไม่มีการแบ่งตัวแปรออกเป็นตัวแปรอิสระและตัวแปรตาม และมักใช้เป็นวิธีเริ่มต้นเพื่อลดจำนวนตัวแปรให้เหลือน้อยลง แล้วนำตัวแปรใหม่ที่สร้างขึ้นไปเป็นตัวแปรอิสระเพื่อทำการวิเคราะห์ด้วยวิธีการอื่นๆต่อไป เช่น การวิเคราะห์ความถดถอยและสหสัมพันธ์ (Regression and Correlation Analysis) การวิเคราะห์ความแปรปรวน (ANOVA) และการวิเคราะห์จำแนกกลุ่ม (Discriminant Analysis) เป็นต้น

### 1) วัตถุประสงค์

- เพื่อลดจำนวนตัวแปรให้เหลือน้อยลงด้วยการจับกลุ่มของตัวแปรที่มีความสัมพันธ์กันไว้ในปัจจัยเดียวกัน เพื่อเป็นตัวแทนของตัวแปรหลายๆตัว
- เพื่อกำหนดจำนวนของปัจจัยที่มีอิทธิพลเพียงพอที่จะอธิบายกลุ่มของตัวแปรเดิมที่ถูกนำมาพิจารณา
- เพื่อตรวจสอบความสัมพันธ์ของแต่ละปัจจัยและความสัมพันธ์ของแต่ละตัวแปร

### 2) หลักการของวิเคราะห์ปัจจัย

การวิเคราะห์ปัจจัยมีข้อสมมติฐานเบื้องต้นว่าเวกเตอร์ตัวแปร  $X$  สามารถเขียนเป็นฟังก์ชันของปัจจัยร่วม (common factor)  $F$  และค่าเฉพาะ (unique factor)  $E$  ของตัวแปรแต่ละตัว โดยที่ปัจจัยร่วมจะแสดงถึงความร่วมกันของตัวแปรหลายตัว หรือปัจจัยร่วมเป็นปัจจัยที่ประกอบด้วยรายละเอียดหรือความแปรผันของตัวแปรหลายตัว

ในกรณีที่มีตัวแปร  $p$  ตัว คือ  $X_1, X_2, \dots, X_p$  ปัจจัยร่วม  $m$  ปัจจัย คือ  $F_1, F_2, \dots, F_m$  และค่าเฉพาะ  $p$  ตัว คือ  $E_1, E_2, \dots, E_p$   
 ตัวแบบ (model) ของการวิเคราะห์ปัจจัยคือ

$$\underline{X}_{p \times 1} - \underline{\mu}_{p \times 1} = \underline{L}_{p \times m} \underline{F}_{m \times 1} + \underline{\epsilon}_{p \times 1} \quad (2.11)$$

### 3) ข้อสมมติเบื้องต้นของการวิเคราะห์ปัจจัย (Assumption)

ข้อสมมติเบื้องต้นสำหรับตัวแบบเชิงตั้งฉากของการวิเคราะห์ปัจจัยมีดังนี้

- ปัจจัยร่วม  $F_i$  และ  $F_j$  ไม่มีความสัมพันธ์กัน
- ค่าเฉพาะของ  $\epsilon_i$  และ  $\epsilon_j$  ไม่มีความสัมพันธ์กัน
- ปัจจัยร่วม  $F_i$  และค่าเฉพาะของ  $\epsilon_j$  ไม่มีความสัมพันธ์กัน

### 4) ขั้นตอนวิเคราะห์ปัจจัย

สำหรับงานวิจัยที่มีตัวแปรที่เกี่ยวข้องหลายๆตัว นักวิจัยต้องการที่จะลดตัวแปร เพื่อสร้างตัวแปรใหม่ ซึ่งเป็นปัจจัยร่วมของตัวแปรหลายๆตัว ขั้นตอนการวิเคราะห์หรือสร้างปัจจัยร่วมมีดังนี้

- ขั้นที่ 1 สร้างเมทริกซ์สัมประสิทธิ์สหสัมพันธ์ของตัวแปรทุกคู่ (correlation matrix)

ถ้าค่าสัมประสิทธิ์สหสัมพันธ์ของตัวแปรคู่ใดมีค่าเข้าใกล้ +1 หรือ -1 แสดงว่าตัวแปรคู่นั้นมีความสัมพันธ์กันมากควรอยู่ในปัจจัยเดียวกัน ถ้าค่าสัมประสิทธิ์สหสัมพันธ์ของตัวแปรคู่ใดมีค่าใกล้ศูนย์ แสดงว่าตัวแปรคู่นั้นไม่มีความสัมพันธ์กันหรือสัมพันธ์กันน้อยควรอยู่คนละปัจจัย ถ้ามีตัวแปรที่ไม่มีความสัมพันธ์กับตัวแปรอื่น ๆ หรือมีความสัมพันธ์กับตัวแปรอื่น ๆ น้อยมาก ให้ตัดตัวแปรนั้นออกจากการวิเคราะห์

- ขั้นที่ 2 การสกัดปัจจัย (factor extraction)

หลังจากที่ตรวจสอบว่าตัวแปรชุดนั้นเหมาะสมที่จะใช้การวิเคราะห์ปัจจัยแล้ว ขั้นตอนต่อไปคือการสกัดปัจจัย เป็นการสกัดหรือการนำความแปรผันของตัวแปรต่างๆไปไว้ในปัจจัยร่วม เพื่อให้ปัจจัยร่วมเป็นตัวแทนของตัวแปรต่างๆ วิธีการสกัดปัจจัยมีหลายวิธี วิธีที่นิยมกันมากที่สุดคือ การวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis หรือ PCA)

การวิเคราะห์องค์ประกอบหลักหรือ PCA) เป็นเทคนิคที่มีวัตถุประสงค์ที่จะนำรายละเอียดของตัวแปรที่มีจำนวนตัวแปรมาก ๆ มาไว้ในองค์ประกอบที่มีเพียงไม่กี่ตัว โดยจะพิจารณาจากรายละเอียดทั้งหมดจากแต่ละตัวแปร ในการวิเคราะห์ PCA จะสร้างการเชื่อมรวมกันเชิงเส้น (linear combination) ของตัวแปร โดยที่

ปัจจัยที่ 1 จะเป็นการรวมกันเชิงเส้นของตัวแปรที่มีรายละเอียดจากตัวแปรทั้งหมดมากที่สุด หรือกล่าวได้ว่ามีค่าแปรปรวนมากที่สุด

ปัจจัยที่ 2 จะเป็นการรวมกันเชิงเส้นของตัวแปร และนำรายละเอียดที่เหลืออยู่มากที่สุดจากตัวแปร โดยที่ปัจจัยที่ 2 จะต้องตั้งฉาก (orthogonal) กับปัจจัยที่ 1 หรือปัจจัยที่ 2 ไม่มีความสัมพันธ์กับปัจจัยที่ 1

ปัจจัยที่ 3 จะเป็นการรวมกันเชิงเส้นของตัวแปร และนำรายละเอียดที่เหลืออยู่มากที่สุดจากตัวแปร โดยที่ปัจจัยที่ 3 จะต้องตั้งฉากกับปัจจัยที่ 1 และ 2 หรือปัจจัยที่ 3 ไม่มีความสัมพันธ์กับ ปัจจัยที่ 1 และ 2

เป็นเช่นนี้ไปเรื่อยๆจนถึงปัจจัยสุดท้าย

ในขั้นตอนนี้สามารถประมาณค่าถ่วงน้ำหนักปัจจัย (factor loading) ได้ แล้วใช้ค่าถ่วงน้ำหนักปัจจัยในการพิจารณาว่ามีตัวแปรใดบ้างที่ควรอยู่ในปัจจัยเดียวกัน โดยในแต่ละปัจจัยให้พิจารณาค่าถ่วงน้ำหนักปัจจัยของแต่ละตัวแปร ถ้าค่าถ่วงน้ำหนักปัจจัยของตัวแปรใดมีค่ามาก (เข้าใกล้ +1 หรือ -1) เราควรจัดตัวแปรนั้นให้อยู่ในปัจจัยดังกล่าว

- ขั้นที่ 3 หาจำนวนปัจจัยร่วมที่เหมาะสม (number of common factors)

จากการสกัดปัจจัยเพื่อลดจำนวนตัวแปรให้เหลือน้อยลง และให้ปัจจัยร่วมเป็นตัวแทนของตัวแปร โดยที่ปัจจัยร่วมจะมีจำนวนน้อยกว่าจำนวนตัวแปร ปัญหาคือควรมีปัจจัยร่วมกี่ปัจจัยจึงจะเหมาะสม ซึ่งทำได้หลายวิธี เช่น พิจารณาเฉพาะปัจจัยร่วมที่มีค่าไอเกนมากกว่า 1 หรือใช้พล็อตกราฟ Scree plot ซึ่งในทางปฏิบัติมักจะพิจารณาจากหลายๆวิธีร่วมกัน

- ขั้นที่ 4 ให้ความหมายปัจจัยร่วม (factor meaning)

จากขั้นที่ 3 เมื่อทราบจำนวนปัจจัยร่วม และคำนวณค่าสัมประสิทธิ์หรือค่าถ่วงน้ำหนักปัจจัยจากขั้นที่ 2 แล้วทำให้ทราบว่าตัวแปรใดที่มีส่วนร่วมในปัจจัยร่วมมาก ก็จะทำให้ความหมายของปัจจัยร่วมเป็นความหมายของกลุ่มตัวแปรที่มีค่าถ่วงน้ำหนักปัจจัยในปัจจัยร่วมนั้นมาก

บางกรณีอาจมีค่าถ่วงน้ำหนักปัจจัยของบางตัวแปรมีค่าต่ำในทุกๆปัจจัยร่วม หรือบางตัวแปรอาจมีค่าถ่วงน้ำหนักปัจจัยปานกลางในทุกๆปัจจัยร่วม ทำให้ไม่สามารถจัดได้ชัดเจนว่าตัวแปรใดบ้างที่มีส่วนร่วมในปัจจัยร่วมแต่ปัจจัย จะต้องทำการหมุนแกนในขั้นที่ 5 ต่อไป

- ขั้นที่ 5 หมุนแกนปัจจัยร่วม (factor rotation)

การหมุนแกนปัจจัยร่วมทำขึ้นเพื่อแก้ปัญหาในกรณีที่เกิดความไม่ชัดเจนในการจัดตัวแปรของปัจจัยร่วมหรือในการให้ความหมายของปัจจัยร่วม

วัตถุประสงค์ของการหมุนแกนปัจจัยคือ เพื่อทำให้ค่าถ่วงน้ำหนักของตัวแปร มีค่ามากขึ้นหรือลดลงจนทำให้ทราบว่าตัวแปรนั้นควรอยู่ในปัจจัยใดหรือไม่ควรอยู่ในปัจจัยใด การหมุนแกนมี 2 วิธี คือ



(1) การหมุนแกนปัจจัยร่วมแบบตั้งฉากกัน (Orthogonal Rotation) เป็นการหมุนแกนไปแล้วยังคงทำให้ปัจจัยตั้งฉากกันหรือเป็นอิสระกัน จะทำให้ปัจจัยร่วมไม่มีความสัมพันธ์ วิธีการหมุนแกนที่นิยมใช้กันมากที่สุด คือ Varimax

(2) การหมุนแกนปัจจัยร่วมแบบเป็นมุมเฉียง (Oblique Rotation) เป็นการหมุนแกนไปในลักษณะที่ปัจจัยไม่ตั้งฉากกัน หรือไม่เป็นอิสระกัน จะทำให้ปัจจัยร่วมมีความสัมพันธ์

- ขั้นที่ 6 คำนวณค่าคะแนนปัจจัย (factor score)

จากขั้นที่ 5 เมื่อหมุนแกนแล้ว ทำให้สามารถจัดตัวแปรหรือให้ความหมายแก่ปัจจัยร่วมแต่ละปัจจัยได้ ทำให้ปัจจัยร่วมที่ได้เป็นเหมือนตัวแปรที่สร้างขึ้นใหม่ จึงควรคำนวณค่าตัวแปรใหม่เพื่อนำไปใช้ในการวิเคราะห์วิธีอื่นๆต่อไป ค่าปัจจัยร่วมเรียกว่าคะแนน โดยแทนค่าตัวแปรลงในสมการที่ 2.12 [46]ดังนี้

$$F_{ij} = b_1 X_{i1} + b_2 X_{i2} + \dots + b_p X_{ip} \quad (2.12)$$

$i=1,2,\dots,n, j=1,2,\dots,m$

## 2.9 เทคนิคการจำแนกกลุ่มในงานวิจัย

### 2.9.1 การวิเคราะห์จำแนกกลุ่มด้วยเทคนิค Discriminant Analysis

การวิเคราะห์จำแนกกลุ่มด้วยเทคนิค Discriminant Analysis [46]เป็นวิธีการทางสถิติที่ใช้วิเคราะห์จำแนกกลุ่มตั้งแต่ 2 กลุ่มขึ้นไป ด้วยการวิเคราะห์จากตัวแปรตาม 1 ตัว และตัวแปรอิสระตั้งแต่ 1 ตัวขึ้นไป การวิเคราะห์ด้วยวิธีนี้นอกจากจะสามารถจำแนกความแตกต่างระหว่างกลุ่มได้แล้ว ยังสามารถบอกธรรมชาติบางอย่างของการจำแนกกลุ่มได้ด้วย เช่น บอกได้ว่าตัวแปรใดจำแนกได้ดีมากน้อยกว่ากัน นั่นคือ สามารถบอกประสิทธิภาพ หรือนำหนักในการจำแนก ของตัวแปรเหล่านั้น การวิเคราะห์จำแนกกลุ่มเป็นการใช้ตัวแปรพยากรณ์หรือตัวแปรอิสระ ที่ร่วมกันพยากรณ์ตัวแปรตาม ซึ่งเป็นเทคนิคทางสถิติที่คล้ายคลึงกับการวิเคราะห์ถดถอยพหุคูณ (Multiple Regression Analysis)

#### 1) วัตถุประสงค์ของการวิเคราะห์จำแนกกลุ่ม

- เพื่อศึกษาปัจจัยหรือตัวแปรที่ทำให้กลุ่มมีความแตกต่างกัน ซึ่งเป็นการศึกษาว่าตัวแปรอิสระใดบ้างเป็นตัวแปรที่ทำให้กลุ่มต่างกัน โดยสร้างฟังก์ชันการจำแนกกลุ่มที่อยู่ในรูปเชิงเส้น

- เพื่อสร้างสมการจำแนกกลุ่มที่ดีที่สุดจากข้อมูลใช้พยากรณ์หน่วยวิเคราะห์ใหม่ที่ยังไม่ทราบกลุ่มมาก่อนว่าควรอยู่ในกลุ่มใดในอนาคต

## 2) ข้อตกลงเบื้องต้นของการวิเคราะห์จำแนกกลุ่มไว้ 4 ประการดังนี้

- ตัวแปรอิสระมีการแจกแจงปกติหลายตัวแปร (Normality of Independent Variables)

- เมตริกซ์ความแปรปรวนร่วมของตัวแปรอิสระของกลุ่มตัวอย่างต้องเท่ากัน (Equal Dispersion Matrices)

## 3) วิธีการตรวจสอบข้อตกลงเบื้องต้น

- ตัวแปรอิสระมีการแจกแจงปกติหลายตัวแปร (Normality of Independent Variables) ตรวจสอบได้จากการแจกแจงปกติแบบตัวแปรเดียว (Univariate Normal Distribution) โดยแยกทำการตรวจสอบตัวแปรอิสระทีละตัว หากพบว่าตัวแปรอิสระทุกตัวมีการแจกแจงปกติ ก็มีความน่าจะเป็นสูงที่จะมีการแจกแจงแบบปกติหลายตัวแปรด้วย

- เมตริกซ์ความแปรปรวนร่วมของตัวแปรอิสระของกลุ่มตัวอย่าง ต้องเท่ากัน (Equal Dispersion Matrices) ตรวจสอบได้ด้วยสถิติ Box's M

## 4) สถิติสำคัญของการวิเคราะห์จำแนกกลุ่ม

- ค่าไอเก้น (Eigenvalue) เป็นค่าที่แสดงอัตราส่วนการผันแปรระหว่างกลุ่มต่อการผันแปรภายในกลุ่ม ถ้าค่าไอเก้นมีค่าสูงก็แสดงว่าสมการดีหรือมีค่าจำแนกสูงหรือกล่าวอีกนัยหนึ่งได้ว่า Eigenvalue ก็คือ Variance ของคะแนนแปลงรูป  $Y$  ที่แปลงมาจาก  $X_1, X_2, \dots, X_p$

- ค่าสหสัมพันธ์คาโนนิคอล (Canonical Correlation) เป็นสถิติซึ่งสามารถใช้ในการตัดสินความสำคัญของสมการจำแนกเป็นมาตรวัดความสัมพันธ์ของสมการกับกลุ่มของตัวแปร ซึ่งระบุการเป็นสมาชิกของกลุ่มนั้น ๆ ของตัวแปรตาม โดยชี้ให้เห็นว่าการเป็นสมาชิกกลุ่มมีความสัมพันธ์กับสมการที่หามาได้มากน้อยเพียงใด ดังนั้น ถ้าค่าสหสัมพันธ์คาโนนิคอลมีค่าสูง แสดงว่าการเป็นสมาชิกของกลุ่มสามารถอธิบายความผันแปรของตัวแปรกับสมการจำแนกได้มาก

- ค่าวิลค์แลมบ์ดา (Wilks' Lambda) เป็นสถิติที่ใช้ทดสอบความแตกต่างระหว่างกลุ่ม และเป็นมาตรวัดอำนาจในการจำแนกกลุ่มของตัวแปรด้วย ถ้าค่าวิลค์แลมบ์ดามีค่ามาก ตัวแปรจะอธิบายการเป็นสมาชิกของกลุ่มได้น้อยถ้าค่าวิลค์แลมบ์ดา มีค่าน้อย ตัวแปรจะอธิบายการเป็นสมาชิกของกลุ่มได้มาก

## 5) วิธีการวิเคราะห์จำแนกกลุ่ม

- คำนวณค่า Eigenvalue

ในการวิเคราะห์จำแนกประเภท จะต้องหาค่า Eigenvalue หรือที่เรียกว่า Discriminant Criterion หรือ Characteristic roots หรือ Latent roots เขียนแทนด้วย สัญลักษณ์  $\lambda$

Eigenvalue ( $\lambda$ ) คือความแปรปรวนของคะแนนแปลงรูป Y ที่แปลงมาจาก  $X_1, X_2, \dots, X_p$  เขียนในรูปสูตรได้เป็น

$$\lambda = \frac{SS_b(Y)}{SS_w(Y)} \quad (2.13)$$

เมื่อ  $SS_b(Y)$  แทน Sum of Square of between group จากคะแนน Y

$SS_w(Y)$  แทน Sum of Square of within group จากคะแนน Y

ค่า  $\lambda$  อาจมีได้หลายค่า จำนวนของค่า  $\lambda$  จะเท่ากับจำนวนกลุ่มลบด้วย 1 (k-1) หรือเท่ากับจำนวนตัวแปร (p) แล้วแต่ว่าจำนวนใดจะน้อยกว่ากัน

- คำนวณหาค่า V แต่ละชุด

หลังจากที่คำนวณค่า  $\lambda$  แต่ละค่าแล้ว นำค่าเหล่านี้ไปคำนวณหาค่า V แต่ละชุด ค่า  $\lambda_1$  จะให้ค่า  $V_1$  ค่า  $\lambda_2$  จะให้ค่า  $V_2$  และค่า  $\lambda_3$  ก็จะให้ค่า  $V_3$  ดำเนินการจนครบ  $\lambda$  ทุกค่า แต่ละค่าดำเนินการดังนี้จาก

$$(W^{-1}B - \lambda_i I)V = 0$$

ขั้นที่ 1 นำเอาค่า  $\lambda$  ไปแทนค่าใน  $(W^{-1}B - \lambda_i I)$  และคำนวณออกมา

ขั้นที่ 2 คำนวณ  $\text{adj}(W^{-1}B - \lambda_i I)$

ขั้นที่ 3 นำเอาค่าในคอลัมน์ใดคอลัมน์หนึ่งของ  $\text{adj}(W^{-1}B - \lambda_i I)$  มายกกำลังสอง รวมกันและถอดรากที่สอง นำไปหารค่าเดิมแต่ละค่าผลที่ได้จะเป็น V ที่สอดคล้องกับ  $V_1$  นั้น

กรณีที่ผู้วิจัยต้องการเปรียบเทียบประสิทธิภาพในการจำแนกของแต่ละตัวแปร ก็ควรแปลง  $V_1$  แต่ละตัวในเมตริกซ์ V ให้เป็นรูปมาตรฐานคือ  $V_{mi}^*$  โดยใช้สูตร

$$V_{mi}^* = \sqrt{W_{ii}} V_{mi} \quad (2.14)$$

เมื่อ  $V_{mi}$  แทน น้ำหนักของการจำแนกที่เป็นมาตรฐานของสมการจำแนกกลุ่ม

$W_{ii}$  แทน สมาชิกในแนวทแยงของเมตริกซ์ W (เมตริกซ์ของผลรวมของกำลังสองและของผลคูณภายในกลุ่ม)

$V_{mi}$  แทน น้ำหนักของการจำแนกในรูปคะแนนดิบของสมการจำแนกกลุ่ม

การเปรียบเทียบค่า V เพื่อบ่งชี้ถึงประสิทธิภาพในการจำแนก และการบ่งชี้ประสิทธิภาพการจำแนกจะมีประสิทธิภาพมากขึ้นเมื่อเราเปรียบเทียบในรูปของคะแนนมาตรฐาน ซึ่งจะจัดอิทธิพลจากหน่วยที่ใช้ในตัวแปรนั้น ๆ ได้

- เขียนสมการจำแนก (Discriminant Function)

เขียนสมการจำแนกโดยการนำเอาค่า V แต่ละชุดมาเขียนสมการจำแนกกลุ่ม โดยมี รูปสมการดังนี้

$$\begin{aligned}
 Y_1 &= V_{11}X_1 + V_{12}X_2 + \dots + V_{1p}X_p \\
 Y_2 &= V_{21}X_1 + V_{22}X_2 + \dots + V_{2p}X_p \\
 &\dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \\
 Y_k &= V_{k1}X_1 + V_{k2}X_2 + \dots + V_{kp}X_p
 \end{aligned}
 \tag{2.15}$$

- การทดสอบนัยสำคัญ

ขั้นตอนนี้เป็นขั้นตอนการทดสอบนัยสำคัญของสมการจำแนกกลุ่ม เมื่อได้สมการจำแนก (Discriminant function) จากขั้นตอนที่ 3 แล้ว ก็จะทำทดสอบนัยสำคัญของสมการที่ได้ เพื่อทราบว่าสมการเหล่านั้นสมการใดมีอำนาจจำแนกกลุ่มได้อย่างมีนัยสำคัญ โดยใช้วิธีของ Barlett test จากสูตร

$$V_m = [N-1- .5(p+k)]\ln(1+ \lambda_m) \tag{2.16}$$

เมื่อ  $V_m$  แทน ค่าสถิติที่จะใช้เปรียบเทียบกับค่าวิกฤตเพื่อทราบความมีนัยสำคัญของสมการที่ m ค่าวิกฤต (Critical Value) หาได้จากการเปิดตาราง ไควสแควร์ที่  $df=p+k - 2m$

- N แทน จำนวนสมาชิกในกลุ่มตัวอย่างทั้งหมด
- P แทน จำนวนตัวแปร
- K แทน จำนวนกลุ่ม
- $\lambda$  แทน Eigenvalue ของสมการที่ทดสอบ

สมการจำแนกจะมีนัยสำคัญ เมื่อค่า  $V_m$  ที่คำนวณได้มีค่ามากกว่าหรือเท่ากับวิกฤต (Critical Value)

ในกรณีที่ต้องการทราบว่าสมการจำแนกกลุ่มรวมกันแล้วสามารถจำแนกกลุ่มได้ อย่างมีนัยสำคัญหรือไม่ ทดสอบได้จากสูตร

$$v=N-1- .5(p+k) \sum_{m=1}^r \ln(1+\lambda_m) \tag{2.17}$$

เมื่อ r แทน จำนวนสมการ

df แทน ชั้นแห่งความเป็นอิสระ ในกรณีนี้มีค่าเท่ากับ  $p(k-1)$

ในกรณีที่ผู้วิจัยทำการทดสอบนัยสำคัญของแต่ละสมการ ก็จะสามารถคำนวณหา ค่า  $V_m$  ได้ถ้านำค่า  $V_m$  ของแต่ละสมการมารวมกันก็ได้ค่า  $V$  ซึ่งเป็นค่าที่ใช้ทดสอบสมการจำแนก กลุ่มโดยภาพรวม

สมการจำแนกกลุ่มแต่ละสมการจะไม่สัมพันธ์กัน นั่นคือ  $Y_1, Y_2, \dots, Y_p$  เป็นอิสระจากกัน สมการจำแนกกลุ่มสมการหลัง ๆ มักส่งผลน้อยมากจนบางครั้งไม่จำเป็นต้องนำมาพิจารณา หรือเสนอสมการในรายงาน ซึ่งอาจพิจารณาได้จากการทดสอบนัยสำคัญของสมการ ถ้าพบว่าไม่มีนัยสำคัญแสดงว่าสมการนั้นส่งผลน้อยมาก

สมการจำแนกกลุ่มสมการแรก ( $V_1$ ) มีอำนาจจำแนกสูงสุด สมการต่อมา มีอำนาจ จำแนกรองลงมาตามลำดับ เมื่อต้องการเปรียบเทียบว่าสมการจำแนกกลุ่มแต่ละสมการ

มีสัดส่วนจำแนกได้มากน้อยอย่างไร สามารถพิจารณาได้จากสูตร

$$p_i = \frac{\lambda_i}{(\lambda_1 + \lambda_2 + \dots + \lambda_m)} \quad (2.18)$$

เมื่อ  $p$  แทน ค่าอำนาจในการจำแนกตัวแปร

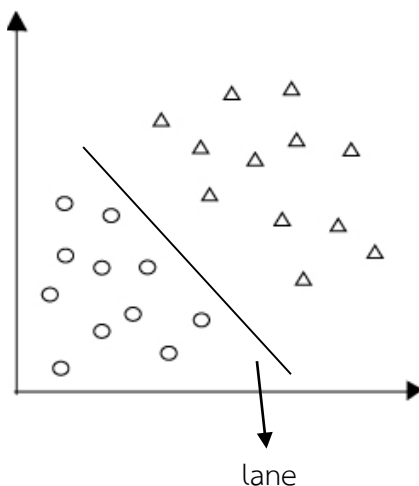
$\lambda$  แทน ค่า Eigenvalue

## 2.9.2 ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine)

ซัพพอร์ตเวกเตอร์แมชชีน[47] เป็นเทคนิคที่ใช้ในการแก้ปัญหาทางด้านการรู้จำรูปแบบข้อมูล โดยอาศัยหลักการของการหาสัมประสิทธิ์ของสมการเพื่อสร้างเส้นแบ่งแยกกลุ่มข้อมูลที่ถูกต้องเข้าสู่กระบวนการสอนให้ระบบเรียนรู้ โดยเน้นไปยังเส้นแบ่งแยกแยะกลุ่มข้อมูลได้ดีที่สุด (Optimal Separating Hyperplane) แต่เดิมซัพพอร์ตเวกเตอร์แมชชีน ถูกนำมาใช้กับข้อมูลที่เป็นเชิงเส้น แต่ในความเป็นจริงแล้วข้อมูลที่นำมาใช้ในระบบการสอนให้ระบบเรียนรู้ส่วนใหญ่มักเป็นข้อมูลแบบไม่เป็นเชิงเส้น ซึ่งสามารถแก้ปัญหาดังกล่าวด้วยการนำเคอร์เนลฟังก์ชันมาใช้

### 1) แนวคิดซัพพอร์ตเวกเตอร์แมชชีน

แนวความคิดของซัพพอร์ตเวกเตอร์แมชชีนนั้น เกิดจากการที่นำค่าของกลุ่มข้อมูลมาวางลงในฟีเจอร์สเปซ (feature space) จากนั้นจึงหาเส้นที่ใช้แบ่งข้อมูลทั้งสองออกจากกัน โดยจะสร้างเส้นแบ่ง (plane) ที่เป็นเส้นตรงขึ้นมา และเพื่อให้ทราบว่าเส้นตรงที่แบ่งสองกลุ่มออกจากกันนั้น เส้นตรงใดเป็นเส้นที่ดีที่สุด ดังรูปที่ 2.19



รูปที่ 2.19 ตำแหน่งข้อมูลสองกลุ่มในพีเจอร์สเปซ

สามารถนำมาเขียนเป็นสมการเพื่อใช้ในการแก้ปัญหา โดยข้อมูลที่นำมาวางลงในพื้นที่คุณลักษณะนั้นเป็นกลุ่มข้อมูลที่อยู่ในรูปของเวกเตอร์จะได้

$$X = ((x_1, y_1)), \dots, ((x_i, y_i)) \quad (2.19)$$

เมื่อ  $X$  คือ ชุดค่าลักษณะเด่น

จากนั้นทำการกำหนดสมการขึ้นมา ซึ่งสมการดังกล่าวมีลักษณะเป็นสมการเส้นตรง เพื่อนำมาสร้างเป็นเส้นตรงบนไฮเปอร์เพลนซึ่งแบ่งกลุ่มข้อมูลที่มีลักษณะเชิงเส้นสองกลุ่มออกจากกัน โดยมีการกำหนดกลุ่มของข้อมูลทั้งสองฝั่งเป็นเพียงสองค่าซึ่งแทนด้วยค่า  $y$  ซึ่งข้อมูลที่เป็นตัวกำหนดความชันและระนาบที่เกิดขึ้นบนไฮเปอร์เพลนเกิดจากคู่ข้อมูล  $(w, b)$  และกำหนดสมการที่เป็นตัวบ่งบอกข้อมูลแต่ละกลุ่มว่าอยู่ส่วนไหนของเส้นแบ่งไฮเปอร์เพลนแสดงได้ ดังสมการที่ 2.20 และสมการที่ 2.21

เมื่อนำสมการเงื่อนไขทั้งหมดมาวิเคราะห์เชิงเรขาคณิต โดยพิจารณาในกรณีข้อมูลที่ถูกแบ่งกลุ่มได้สมบูรณ์ตามเงื่อนไขของซัพพอร์ตเวกเตอร์แมชชีนในสมการที่ 2.22 และข้อมูลการสอนให้ระบบเรียนรู้ต้องอยู่ในรูปแบบของเชิงเส้น เวกเตอร์ของข้อมูลที่ถูกป้อนเข้าสู่ระบบการสอน เพื่อให้ระบบเรียนรู้แทนด้วยสมการ และข้อมูลทั้งสองด้านแบ่งเป็นบวกและลบ สถานะของข้อมูลจึงแทนด้วย  $y$  ซึ่งมีสองค่าคือ  $y = 1$  และ  $y = -1$  นั้น แต่ทั้งนี้ก็ยังตัดสินใจไม่ได้ว่าเส้นแบ่งนั้นควรจะเป็นเส้นใดจึงจะดีที่สุด วิธีการที่ใช้ในการหาเส้นแบ่งที่ดีที่สุดคือการเพิ่มเส้นขอบให้กับเส้นแบ่งทั้งสองข้าง ทำให้ได้เส้นใหม่ที่จะถือเป็นเส้นขอบของข้อมูลแต่ละฝั่งอีกด้วย เส้นขอบ (Margin) ที่เป็นเส้น

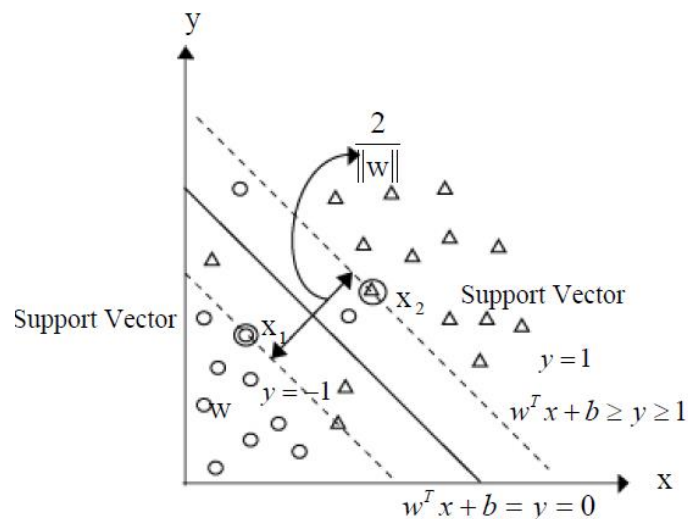
ขอบของเส้นแบ่งนั้นจะเป็นเส้นที่สัมผัสกับค่าข้อมูลในพีเจอรส์เปซที่ใกล้ที่สุด เส้นขอบของทั้งสองเส้นที่เพิ่มขึ้นมานี้ถูกแทนด้วยสมการ  $w^T x + b \geq y \geq 1$  ถ้าอยู่ด้าน  $y = 1$  และ  $w^T x + b \leq y \leq -1$  ถ้า  $y = -1$  หากเส้นขอบของเส้นแบ่งใด ๆ ที่มีความกว้างมากที่สุด แสดงให้เห็นว่าข้อมูลสองชุดมีการแยกกันชัดเจนมากที่สุด ดังนั้นเส้นแบ่งที่มีเส้นขอบกว้างที่สุดจึงเป็นเส้นแบ่งที่ดีที่สุด และเรียกตำแหน่งการสัมผัสข้อมูลที่ใกล้ที่สุดจากการเพิ่มขอบนี้ว่า “เวกเตอร์สนับสนุน” (Support Vector) โดยเรียกเส้นประที่แบ่งข้อมูลทั้งสองเส้นว่าเส้นขอบ ซึ่งสามารถเขียนเป็นสมการการคำนวณ ความกว้างของเส้นขอบต้องทำการคำนวณพจน์ให้อยู่ในรูปปกติ (Normalization) โดยคำนวณจากสมการที่ 2.20 และ 2.21 เมื่อแทนค่าลงไปแล้ว

$$w^T x + b \geq y \quad \text{เมื่อกำหนดให้ } y = 1 \quad (2.20)$$

$$w^T x + b \leq y \quad \text{เมื่อกำหนดให้ } y = -1 \quad (2.21)$$

$$y(w^T x + b) - 1 \geq 0 \quad (2.22)$$

เมื่อ	$y$	คือ ค่ากลุ่มข้อมูล (1, -1)
	$W$	คือ ค่าความชัน $w$
	$X$	คือ ค่าลักษณะเด่น
	$B$	คือ ค่าคงที่ (ค่าตัดแกน $y$ )



รูปที่ 2.20 การวางตัวของข้อมูลในลักษณะเชิงเส้น[48]

$$w^T x^+ + b = 1$$

$$w^T x^- + b = -1$$

$$w^T(x^+ - x^-) = 2$$

$$M = \left(\frac{w}{\|w\|}\right)^T (x^+ - x^-)$$

$$M = \frac{w}{\|w\|} \quad (2.23)$$

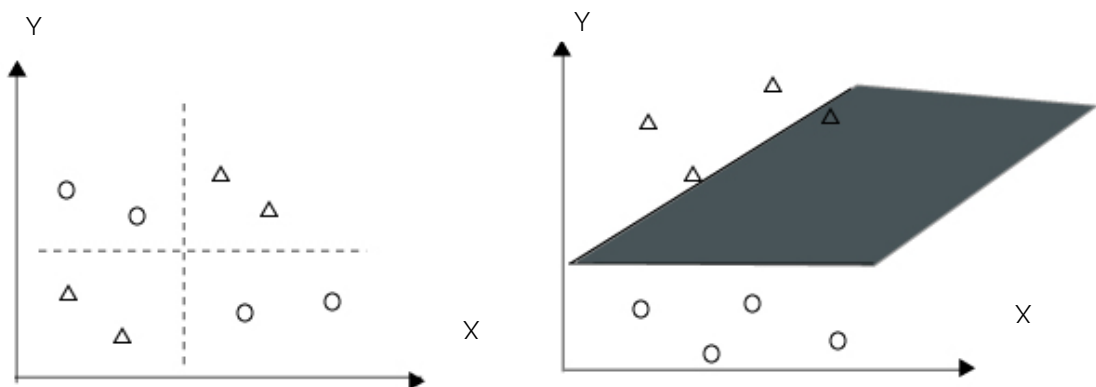
เมื่อ  $M$  คือ ความกว้างของเส้นขอบ

### 3) ข้อผิดพลาดที่ไม่สามารถแบ่งกลุ่มข้อมูลได้

จากการหาซัพพอร์ตเวกเตอร์แมชชีนข้างต้น คิดในกรณีที่ข้อมูลไม่เกิดข้อผิดพลาด แต่ในความเป็นจริงผลลัพธ์ของการแบ่งข้อมูลด้วยซัพพอร์ตเวกเตอร์แมชชีน ย่อมมีข้อผิดพลาดเกิดขึ้นจากค่าข้อมูลที่วางตัวในแกนซัพพอร์ตเวกเตอร์ การแก้ปัญหานี้คือการยอมให้มีการเกิดข้อผิดพลาดนี้ได้ในระยะหนึ่งที่ยอมรับได้ ค่านั้นเรียกว่า “เวกเตอร์อนุโลม” (slack vector) ดังแสดงภาพที่ 2.20

### 4) เคอร์เนล (Kernel)

เมื่อข้อมูลที่ใช้สำหรับการรู้จำมีการวางตัวกันในลักษณะตำแหน่งที่ไม่สามารถแบ่งกลุ่มได้โดยเส้นตรงตามวิธีการของซัพพอร์ตเวกเตอร์แมชชีนที่มีรากฐานมาจากข้อมูลเชิงเส้น จึงไม่สามารถนำมาใช้กับข้อมูลที่ไม่เป็นเชิงเส้นในลักษณะเช่นนี้ได้ โดยสามารถแก้ปัญหาลักษณะเช่นนี้ได้โดยการเปลี่ยนแปลงมิติของข้อมูลให้สูงขึ้นเพื่อช่วยในการเรียงตัวของข้อมูลเสียใหม่ที่เรียกว่า “พื้นที่มิติสูง” (higher dimensional space)



รูปที่ 2.21 การเปลี่ยนมิติข้อมูลไปยังมิติที่สูงขึ้นด้วยฟังก์ชันของเคอร์เนล[48]

ในเคอร์เนลนั้นคือการคูณกันของชุดเวกเตอร์ของ  $x$  ใด ๆ



$$K(x_i, x_j) = x_i^T x_j \quad (2.24)$$

เคอร์เนลที่นิยมใช้มีอยู่ 3 ชนิดด้วยกันคือ

- โพลีโนเมียล (Polynomial):

$$K(x_i, x_j) = (\lambda x_i^T x_j + r)^d, \quad \lambda > 0 \quad (2.25)$$

เมื่อ  $d$  คือ ค่าของ Degree ในรูปเลขยกกำลัง

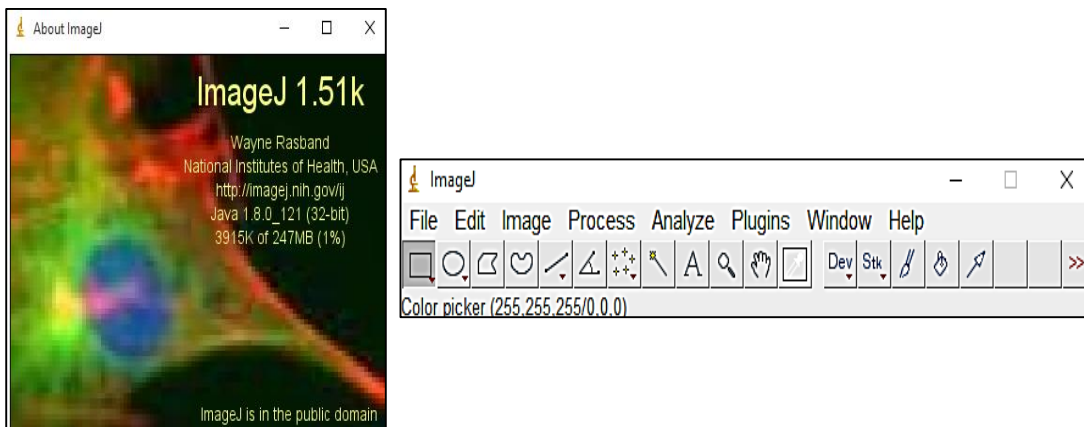
ดังนั้นจากสมการของเคอร์เนลนั้นสามารถที่จะแทนลงไปในตำแหน่งของ  $x_i^T x_j$  ซึ่งสามารถจึงเขียนเป็นสมการได้ดังสมการที่ 2.28 ซึ่งเป็นสมการที่ใช้ในขั้นตอนที่จะเรียนรู้ว่าจะวางตำแหน่งเส้นแบ่งไว้ที่ตำแหน่งใดโดยทำงานร่วมกับเคอร์เนล เพื่อแปลงให้ข้อมูลที่ยากต่อการแบ่งแบบเชิงเส้นสามารถแบ่งได้เมื่อถูกทำให้เป็นข้อมูลแบบมิติสูง (Higher Dimension) ดังนั้นจึงมีอีกสมการหนึ่งที่ใช้ค่า  $w$  และ  $b$  เดิมมาจัดตำแหน่งของข้อมูลเพื่อที่ให้ทราบว่าข้อมูลนั้นเป็นกลุ่มใด กำหนดได้ดังสมการที่ 2.29

$$y_i (\sum_{i=1}^N \alpha_i y_i K(x_i, x_j) + b) - 1 \geq 0 \quad (2.28)$$

$$f(x) = \text{sgn}(\sum_{i=1}^N \alpha_i y_i K(x_i, x_j) + b) \quad (2.29)$$

## 2.10 โปรแกรมที่ใช้ในงานวิจัย

1) โปรแกรม ImageJ[49] เป็นโปรแกรมฟรีแวร์ที่คิดค้นพัฒนาขึ้นเพื่อใช้ในการวิเคราะห์ภาพถ่าย โดยสำนักวิจัยของสถาบัน National Institute of Mental Health (NIMH) ประเทศสหรัฐอเมริกาแสดงในรูปที่ 2.22 ซึ่งใช้ได้กับเครื่องคอมพิวเตอร์ส่วนบุคคล (personal computer) ที่ติดตั้งโปรแกรม Java ตั้งแต่เวอร์ชัน 1.4 ขึ้นไป โดยสามารถเลือกระบบปฏิบัติการได้ เช่น Windows, Linux, Mac OS X เป็นต้น โปรแกรมสามารถเปิดภาพและบันทึกไฟล์นามสกุล TIFF, GIF, JPEG, BMP, DICOM, FITS และ raw และเลือกบันทึกในรูปแบบ 8 บิต, 16 บิต หรือ 32-บิตได้ โปรแกรมนี้เป็นโปรแกรมที่มีประโยชน์อย่างมากในงานวิจัยที่จำเป็นต้องดึงข้อมูลจากภาพถ่าย โดยโปรแกรมสามารถหาพื้นที่ วัดระยะ นับจำนวนของวัตถุบนภาพ และสามารถคำนวณหาสัดส่วนพื้นที่และความยาวได้ในหน่วยการวัดต่างๆ เช่น หน่วยการวัดระบบเมตริก หรือหน่วยพิกเซล (pixel) ของภาพ เป็นต้น นอกจากนี้ยังสามารถแสดงผลในรูปแบบแผนภูมิหรือฮิสโตแกรม (histogram)



รูปที่ 2.22 โปรแกรม ImageJ เวอร์ชัน 1.51k

2) โปรแกรม Weka[50] (Waikato Environment for Knowledge Analysis) เป็นซอฟต์แวร์สำเร็จประเภทฟรีแวร์พัฒนาขึ้นโดยใช้ภาษาจาวา เหมาะกับการใช้ในเทคนิคด้านการเรียนรู้ด้วยเครื่อง (Machine Learning) และการทำเหมืองข้อมูล (Data Mining) โปรแกรมจะประกอบไปด้วยโมดูลย่อยสำหรับการจัดการข้อมูล สามารถใช้ Graphic User Interface (GUI) และใช้คำสั่งในการให้ประมวลผล และสามารถพัฒนาโปรแกรมได้เป็นเครื่องมือที่ใช้ทำงานในด้านการทำเหมืองข้อมูล ซึ่งมีอัลกอริทึมที่หลากหลาย โดยชุดเครื่องมือมีฟังก์ชันสำหรับการทำงานร่วมกับข้อมูล ได้แก่ Pre-Processing, Classification, Regression, Clustering, Association rules, Selection และ Visualization ดังรูปที่ 2.23



รูปที่ 2.23 โปรแกรม Weka เวอร์ชัน 3.8.1

3) โปรแกรมวิเคราะห์ข้อมูลทางสถิติ SPSS (Statistical Package for the Social Science for Windows)[51] เป็นโปรแกรมวิเคราะห์ และการจัดการข้อมูลโดยใช้สถิติ ประเภทต่างๆการใช้งานโปรแกรมไม่ซับซ้อน ให้ผลลัพธ์ที่รวดเร็ว และถูกต้องแม่นยำ แสดงผลการวิเคราะห์ข้อมูลออกมาในรูปของตาราง หรือแผนภูมิชนิดต่าง ๆ ได้ทั้งแบบ 2 มิติ และ 3 มิติ สามารถนำไปประยุกต์ใช้ในงานวิจัย และเครื่องมือที่ช่วยในการตัดสินใจ

## 2.10 การตรวจสอบเอกสาร

จากการสำรวจงานวิจัยที่ผ่านมาพบว่าได้มีการนำเทคนิคการประมวลผลภาพมาประยุกต์ใช้ในการประเมินลักษณะต่างๆที่หลากหลาย เช่น การแพทย์ เกษตรกรรม อุตสาหกรรม เป็นต้น ทั้งในส่วนการตรวจสอบลักษณะ ปริมาณ หรือคุณภาพเพื่อลดความผิดพลาดจากการประเมินด้วยสายตาของมนุษย์ซึ่งอาจจะไม่มีความแม่นยำและสม่ำเสมอ ทำให้ผลการประเมินมีความไม่คงที่แน่นอน โดยเทคนิคการประมวลผลที่จะนำเข้ามาประยุกต์ใช้จะมีความเฉพาะแต่ละประเภท นอกจากนั้นมีความหลากหลาย และความไม่แน่นอน จึงต้องกำหนดเกณฑ์ในการประเมินโดยใช้คุณลักษณะเป็นตัวชี้วัด ซึ่งคุณลักษณะที่นิยมนำมาใช้พารามิเตอร์ในเทคนิคนี้ได้แก่ พื้นที่ เส้นรอบรูป เส้นผ่านศูนย์กลางเฉลี่ย ความเข้มของสี ความกว้าง และความยาว แต่ละงานวิจัยเมื่อกำหนดพารามิเตอร์ที่เหมาะสมแล้ว จะทำการคำนวณค่าแล้วใช้ค่าพารามิเตอร์เหล่านี้ในการจำแนกหรือตัดสินใจข้อมูลซึ่งระบบจะทำการเปรียบเทียบกับข้อมูลตัวอย่างที่อยู่ในแต่ละกลุ่ม เพื่อให้สอดคล้องกับลักษณะงานวิจัยที่จะดำเนินการ จึงนำเสนอผลการสำรวจงานวิจัยในส่วนของเทคนิคที่ใช้ในการประมวลผลภาพ เทคนิคการจำแนกกลุ่ม และงานวิจัยเกี่ยวข้องกับการตรวจสอบลักษณะของกุ้ง ซึ่งแยกเป็น 3 ส่วนดังต่อไปนี้

### 2.10.1 เทคนิคที่ใช้ในการประมวลผลภาพ

#### 2.10.1.1 การแปลงภาพเป็นภาพระดับเทา

ในกระบวนการประมวลผลภาพเมื่อได้รับภาพตัวอย่างแล้วจะทำการการประมวลผลภาพเบื้องต้น (pre-processing) เพื่อให้สะดวกต่อการวิเคราะห์ ซึ่งการแปลงภาพสีเป็นภาพระดับเทาเป็นวิธีการหนึ่งเพื่อช่วยจำกัดลักษณะทางสีไม่จำเป็นนอกจากการวิเคราะห์ ทำให้สามารถวิเคราะห์ผลได้ง่ายยิ่งขึ้น Eerens และคณะ[52], Jayas และคณะ[53] Du และ Sun[54] จะเน้นให้ความสำคัญกับการแปลงภาพสีให้เป็นระดับเทาให้เป็นขั้นตอนระหว่างการประเมินก่อนนำไป

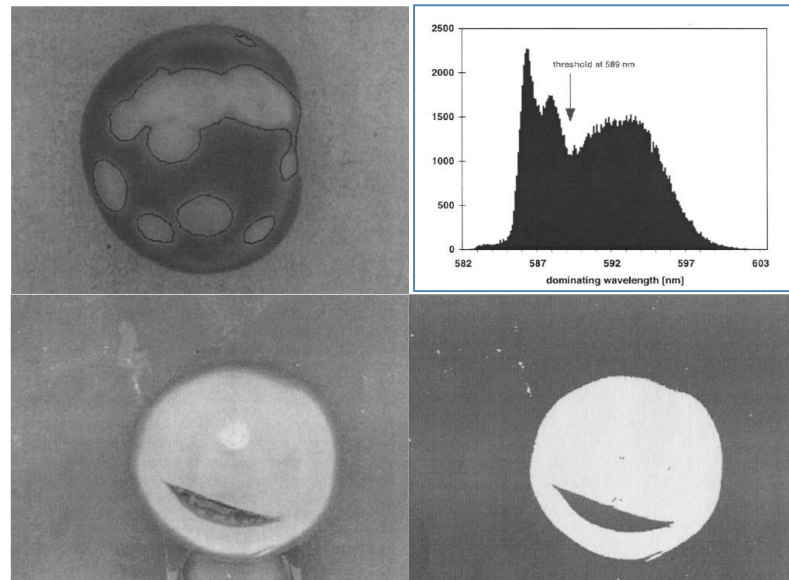
วิเคราะห์คุณลักษณะทางคุณภาพของอาหาร ซึ่งนำมาใช้ประเมินผลในอาหารที่หลากหลายชนิด เช่น ผลไม้, การประมง, ธัญพืช, เนื้อผัก และอื่น ๆ นอกจากนี้ประสิทธิภาพนี้ ประสิทธิ์ นครราช และคณะ[55]ใช้การแปลงภาพเป็นระดับสีเทา เพื่อแยกสายพันธุ์ของข้าวสองสายพันธุ์ ซึ่งเมื่อแปลงภาพแล้วสีของเมล็ดข้าวแต่ละพันธุ์มีความเข้มสีแตกต่างกัน ทำให้สามารถการแบ่งแยกพันธุ์ข้าวได้

### 2.10.1.2 การแยกบริเวณ และคุณลักษณะที่ตรวจสอบ

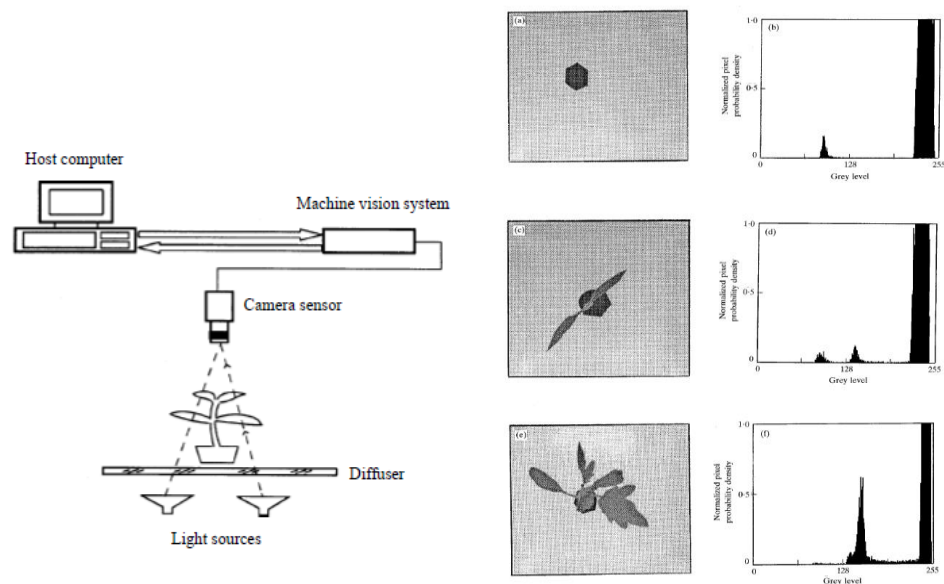
ในงานวิจัยที่มีการแบ่งส่วนประกอบภาพ ใช้ในการแยกองค์ประกอบต่างๆ ของรูปภาพออกจากกันตามลักษณะสำคัญที่พิจารณา ช่วยลดจำนวนข้อมูลในรูปภาพที่ไม่ต้องการหรือไม่จำเป็นออกจากการวิเคราะห์ เนื่องจากในแต่ละภาพมีพารามิเตอร์ในการแบ่งแยกส่วนต่างกัน และแต่ละงานวิจัยจะพิจารณาลักษณะที่ต่างกันขึ้นอยู่กับวัตถุประสงค์ในการประเมิน ในที่นี้จะกล่าวถึงงานวิจัยที่พิจารณาเฉพาะคุณลักษณะด้านสีในการตรวจสอบในงานวิจัยสาขาต่างๆ

งานวิจัยสาขาเกษตรมีการนำเทคนิคการประมวลผลจากภาพมาใช้อย่างกว้างขวางในการตรวจสอบคุณภาพผักผลไม้ชนิดต่างๆ งานวิจัยที่น่าสนใจประกอบด้วย Payman และคณะ[56] ประเมินความเป็นไปได้ของการใช้คอมพิวเตอร์วิทัศน์สำหรับการตรวจสอบคุณภาพแอปเปิลพันธุ์ “Golden Delicious” จากตำหนิบนผิวร่วมกับใช้เทคนิคโครงข่ายประสาทเทียม โดยแยกส่วนก้าน ชั่ว และผลออกจากพื้นหลัง ต่อมา Jahns และคณะ[57]ประเมินคุณภาพสีผิวของมะเขือเทศ ในการกระจายสีของมะเขือเทศสามารถประเมินโดยใช้วิธีแบ่งส่วนภาพ (threshold) ด้วยการพิจารณาลักษณะฮิสโตแกรม(histogram) ดังรูปที่ 2.24 ซึ่งแสดงค่าเกณฑ์ชัดเจนระหว่างจุดสีเหลืองและพื้นที่ผิวสีแดง นอกจากนี้ Ling และ Ruzhitsky[58] ยังนำเทคนิคนี้ไปใช้วัดขนาดต้นมะเขือเทศ โดยบริเวณพุ่มบนของต้นมะเขือเทศมีความสำคัญในการบ่งชี้ขนาด จึงวัดพื้นที่ใบในสภาพแวดล้อมแสงมีการควบคุม แล้วแบ่งส่วนภาพด้วยวิธี Otsu ของต้นกล้ามะเขือเทศเป็นขั้นตอนแรก ดังรูปที่ 2.25

Li และคณะ [59]ศึกษาวิธีการตรวจสอบหารอยตำหนิของส้ม เช่น รอยแมลงกัดรอยที่เกิดจากเพลี้ยไฟทำให้ผิวเปื่อยไหม้ เกิดสะเก็ด และรอยต่าง ซึ่งเปลือกจะมีลักษณะเป็นสีน้ำตาลเข้มขึ้นด้วยวิธีการแบ่งส่วนภาพ (threshold) โดย Caro และคณะ[60]ใช้เทคนิคการวิเคราะห์ภาพถ่ายผลส้มเพื่อจำแนกระดับความสุขของส้มด้วยลักษณะทางกายภาพโดยใช้เกณฑ์มาตรฐาน นอกจากนี้ Wu และคณะ[61] แยกส่วนของเส้นใยจากพื้นหลังด้วยวิธี Otsu ช่วยในการตรวจจับเส้นใยที่ยากต่อมองเห็นได้อย่างรวดเร็วและชัดเจน ทำให้สามารถวิเคราะห์ผลได้แม่นยำมากยิ่งขึ้น



รูปที่ 2.24 มะเขือเทศสีไม่สม่ำเสมอใช้การแบ่งส่วนด้วยฮิสโตแกรม[57]



รูปที่ 2.25 การรับภาพ และแบ่งส่วนภาพของต้นกล้ามะเขือเทศ[58]

### 2.10.1.3 การสกัดคุณลักษณะ

ขั้นตอนการสกัดหรือแยกคุณลักษณะซึ่งเป็นกระบวนการสำคัญที่ดึงคุณลักษณะเฉพาะ เพื่อนำไปเป็นพารามิเตอร์ในการจำแนก และตัดแยกกลุ่ม จากการสำรวจพบว่าในงานวิจัยของ Caro และคณะ[60] ใช้คุณลักษณะของสีผิว ความกว้าง และพื้นที่ของผลเป็นพารามิเตอร์ในการระบุระดับความสุกของส้ม และ Naoshi ร่วมกับคณะ[62] ได้พัฒนาเครื่อง

ตรวจสอบคุณภาพของส้มพันธุ์ “lyokan” ซึ่งใช้คุณลักษณะสี รูปร่าง และความขรุขระของพื้นผิว รวมอัตราส่วนของสี R:G B อัตราเส้นผ่าศูนย์กลาง และเนื้อสัมผัส เป็นพารามิเตอร์วิเคราะห์ด้วยวิธีโครงข่ายประสาทเทียม และงานวิจัยของ Velioğlu และคณะ[63] ทำการหาคุณภาพของมะเขือเทศที่ด้วยการประมวลผลจากภาพที่วัดสีระบบ CIE L\*a\*b\* เพื่อคัดแยกสี นับจำนวน และขนาดของจุดดำ และในงานวิจัยของ Weis และ Gerhards [64] ได้พยายามตรวจจับวัชพืชโดยใช้การแบ่งส่วนภาพเพื่อจำแนกวัชพืชโดยใช้คุณลักษณะทางรูปร่าง ซึ่งสามารถคัดแยกวัชพืชได้ดีที่สุด โดย Ehsanirad และ Kumar[65] ก็ได้นำคุณลักษณะคุณลักษณะดังกล่าวมาใช้ในการจำแนกประเภทของใบพืช Jayas และคณะ[53] ได้ศึกษาลักษณะเชิงพื้นที่ สี และลักษณะพื้นผิวในภาพ เพื่อคัดแยกผลิตภัณฑ์ทางการเกษตร

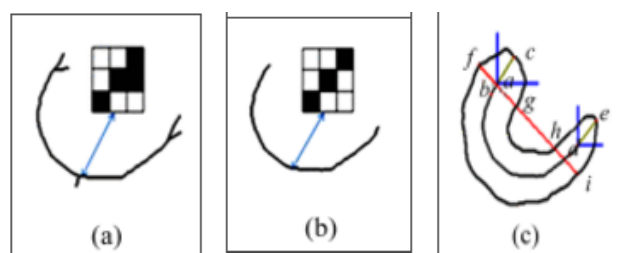
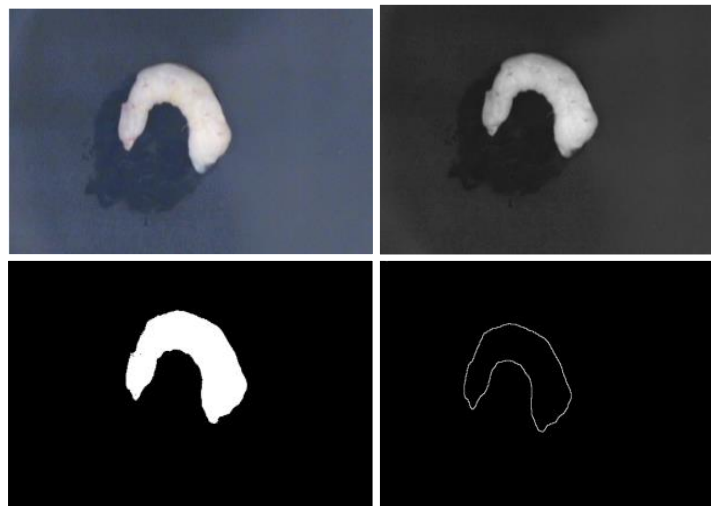
นอกจากนี้ในสาขาของการแพทย์ David Mas และ Belen Ferrer[66] ใช้วิธีการประมวลผลภาพในการตรวจหาเชื้อมาลาเรียเบื้องต้นในหลอดทดลอง โดยใช้การวิเคราะห์พฤติกรรมในเซลล์ที่มีปรสิตมาลาเรียภายในจะมีสีเข้ม จึงใช้พารามิเตอร์ของพื้นที่สีในการวิเคราะห์ ซึ่งผลการทดลองเบื้องต้นมีประสิทธิภาพที่ดีในการจำแนกเซลล์เม็ดเลือดแดงปกติกับเซลล์เม็ดเลือดแดงติดเชื้อปรสิตมาลาเรีย

### 2.10.2 เทคนิคการจำแนกกลุ่ม

หลังจากการประมวลผลภาพเพื่อดึงเอาคุณลักษณะเฉพาะมาวิเคราะห์ด้วยเทคนิคการจำแนกวิธีต่างๆ พบว่า Li และคณะ[59] นำเสนอผลการศึกษาศึกษาการประมวลผลภาพเพื่อจำแนกความสูงอมของส้มซีตรัสด้วยการวิเคราะห์เพื่อแยกส่วนภาพวิธี multispectral และใช้คุณลักษณะของดัชนีสีบนผิวของส้ม แล้วจำแนกด้วย Fisher linear discriminant (FLDA) และงานวิจัยตัวแบบที่ใช้คัดเลือกปัจจัยเด่นจากปัจจัยภายนอกของไซนกระทาเพื่อใช้ในการทำนายเพศของนกระทาโดยจักริน และคณะ[67] สกัดคุณลักษณะได้ 2 ชนิด จาก 7 ชนิด คือความ กว้างและลักษณะด้านแหลมของไซ แล้วทำการวิเคราะห์ผ่านตัวแบบด้วยตัวแบบต้นไม้ในการตัดสินใจ ทฤษฎีของเบย์ โครงข่ายประสาทเทียม ซัพพอร์ตเวกเตอร์แมชชีน ขั้นตอนวิธีการเพื่อนบ้านที่ใกล้ที่สุด และกฎการเชื่อมต่อกัน ซึ่งตัวแบบการทำนายซัพพอร์ตเวกเตอร์แมชชีนให้ค่าความถูกต้องในการแยกเพศสูงที่สุด และงานวิจัยของ Mizushima และคณะ[68] เสนอวิธีการคัดแยกและแบ่งเกรดให้กับแอปเปิ้ล โดยขั้นตอนแรกจะแบ่งส่วนเพื่อแยกเฉพาะส่วนที่สนใจด้วยวิธี Otsu และใช้ตัวอย่างที่มีลักษณะสีแตกต่างกันพัฒนาอัลกอริทึมด้วยซัพพอร์ตเวกเตอร์แมชชีนเชิงเส้น (SVM) ซึ่งได้ผลลัพธ์ที่สม่ำเสมอถูกต้อง และมีข้อผิดพลาดน้อยกว่า 2%

### 2.2.3 งานวิจัยเกี่ยวข้องกับการตรวจสอบลักษณะของกุ้ง

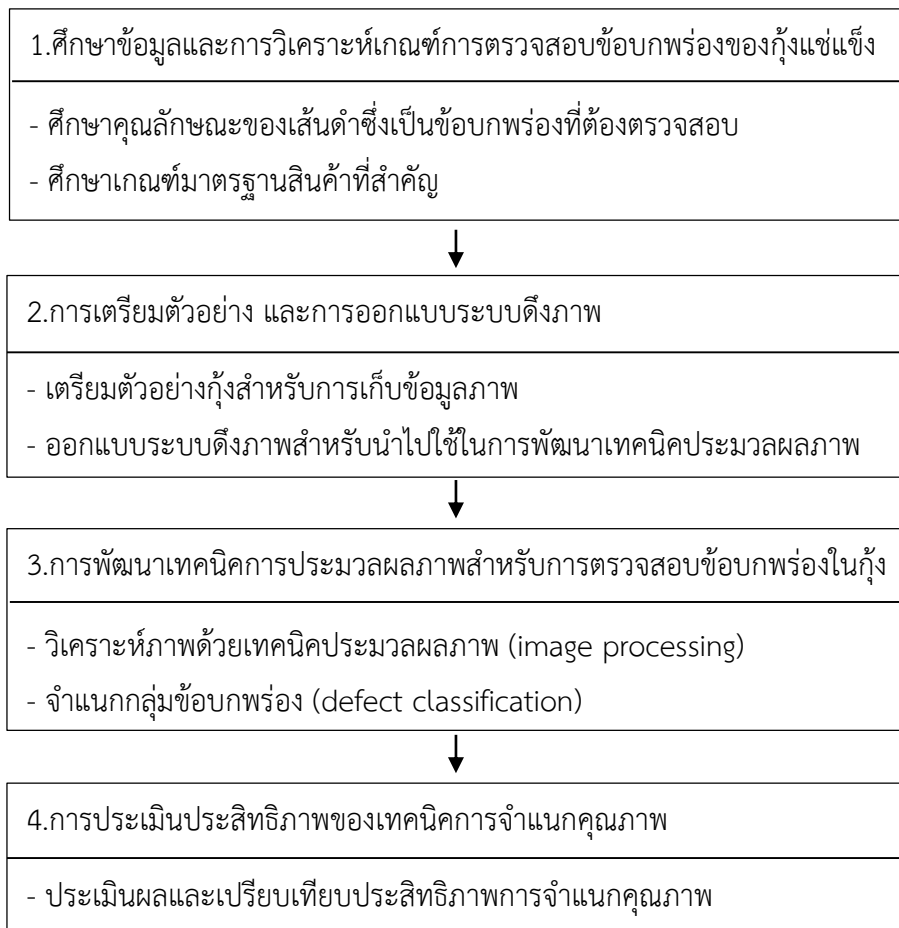
ในการสำรวจผลงานวิจัยที่เกี่ยวกับการใช้การประมวลผลภาพในผลิตภัณฑ์จากกุ้ง พบว่า Jia-Huan Qu และคณะ [69] ใช้เทคนิคการถ่ายภาพ Hyper-spectral ในการคัดแยกลักษณะการลอกเปลือกของกุ้งโดยพิจารณาภาพสะท้อน เพื่อแยกลักษณะกุ้งโดยใช้ตัวแบบเส้นทางกำลังสองน้อยที่สุด (PLS) และวิธีการเรียงลำดับภาพ (SPA) สามารถอธิบายได้มากกว่า 99% นอกจากนี้ประยุกต์ใช้ร่วมกับการเปรียบเทียบด้วยเทคนิคป่าสุ่ม และการจำแนกกลุ่มด้วยวิธี SIMCA งานวิจัยของ M. Mohebbi และคณะ [70] พัฒนาเทคนิคเพื่อใช้ในการประมาณค่าความชื้นในกุ้งแห้ง โดยการประเมินจากคุณสมบัติทางสี  $L^*a^*b$  โดยทำการวิเคราะห์สีระหว่างกระบวนการอบแห้ง เพื่อจำแนกปริมาณความชื้นออกเป็น 4 ระดับ คือ ปริมาณความชื้นระดับสูง ปานกลาง ต่ำ และต่ำมาก ซึ่งผลการศึกษาพบว่า ระบบอัตโนมัติที่ใช้มีความมีประสิทธิภาพมากกว่าวิธีการเดิม และ Peng-min Pan และคณะ [21] ใช้เทคนิคการประมวลผลด้วยภาพหาขนาดของกุ้งที่ลอกเปลือกแล้ว ด้วยการแปลงภาพไบนารี ดังรูปที่ 2.26 หาตัวแปรพื้นที่ ความยาวและความกว้าง วิเคราะห์ด้วยโครงข่ายประสาทเทียม (ANN) ส่งผลให้การทำนายมีความผิดพลาดโดยเฉลี่ยอยู่ที่ 2.67%



รูปที่ 2.26 การแบ่งส่วนและประมวลผลภาพกุ้ง

### บทที่ 3 วิธีการดำเนินงาน

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาเทคนิค (algorithm) การตรวจสอบเส้นดำซึ่งเป็นข้อบกพร่องของกึ่งแข็งด้วยการใช้เทคนิคการประมวลผลภาพ กระบวนการในการวิจัยเป็นการทดลองเพื่อหาคุณลักษณะที่มีนัยสำคัญในการจำแนกคุณภาพตามเกณฑ์มาตรฐานสินค้านำเข้าของประเทศคู่ค้าจากภาพถ่าย แบ่งเนื้อหาออกเป็น 4 ส่วน โดยมีหัวข้อและวิธีดำเนินการดังแผนผังรูปที่ 3.1 และมีรายละเอียดวิธีการดำเนินการวิจัยดังต่อไปนี้



รูปที่ 3.1 หัวข้อและวิธีดำเนินการวิจัย



### 3.1 ศึกษาข้อมูลและการวิเคราะห์เกณฑ์การตรวจสอบข้อบกพร่องของกุ้งแช่แข็ง

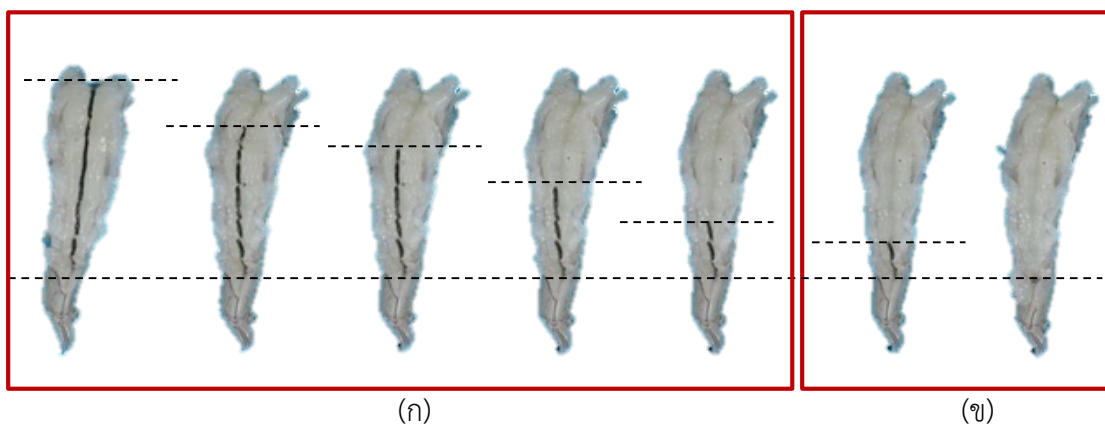
จากการศึกษากระบวนการผลิตกุ้งแช่แข็งในระดับอุตสาหกรรมพบว่า กุ้งจะถูกนำมาแปรรูปเพื่อเพิ่มมูลค่าในหลากหลายผลิตภัณฑ์ โดยรูปแบบการแปรรูปนั้นขึ้นกับความต้องการของประเทศคู่ค้าที่นำเข้าผลิตภัณฑ์ สามารถแบ่งประเภทผลิตภัณฑ์โดยทั่วไปได้ 5 รูปแบบดังตารางที่ 2.1 จากผลิตภัณฑ์ดังกล่าวพบว่าผลิตภัณฑ์ที่ต้องมีการกำจัดเส้นดามี 4 ประเภทด้วยกัน คือ กุ้งสดไม่ลอก เปลือกไฉ่หาง กุ้งลอกเปลือกไฉ่หาง กุ้งสดลอกเปลือกไม่ไฉ่หาง และกุ้งยัด โดยผลิตภัณฑ์กุ้งสดแช่แข็งเป็นผลิตภัณฑ์ที่มีการส่งออกมากที่สุด ดังนั้นงานวิจัยนี้จึงให้ความสนใจ และศึกษาผลิตภัณฑ์กุ้งแช่แข็งที่มีการกำจัดเส้นด้าตามมาตรฐานคุณภาพของประเทศคู่ค้าที่สำคัญอย่างมาตรฐานการจำแนก ระดับกุ้งสดและแช่แข็งประเทศสหรัฐอเมริกา (United States Standards for Grades of Fresh and Frozen Shrimp) ซึ่งเป็นประเทศที่มีสัดส่วนการนำเข้าผลิตภัณฑ์กุ้งมากที่สุด ได้ระบุลักษณะของการกำจัดเส้นด้าที่ไม่เหมาะสมตามหัวข้อ 2.1.4 เกณฑ์มาตรฐานของลักษณะดังกล่าวพิจารณาจากลักษณะความยาวของเส้นด้าที่ปรากฏบนตัวกุ้งเปรียบเทียบกับความยาวของข้อกุ้ง โดยกุ้งที่ไม่ผ่านเกณฑ์มาตรฐานจะมีความยาวรวมทั้งหมดของเส้นด้ามากกว่าหนึ่งข้อลำตัว จากนั้นเทียบน้ำหนักของกุ้งที่ไม่ผ่านเกณฑ์มาตรฐานกับน้ำหนักของกุ้งทั้งหมด และพิจารณาเปรียบเทียบกับตารางข้อบกพร่องสำหรับการแบ่งเกรดของผลิตภัณฑ์กุ้งดังตารางที่ 2.2 ซึ่งเกรดของผลิตภัณฑ์กุ้งที่สามารถยอมรับได้มี 2 เกรด คือ เกรดเอ และเกรดบี วิจัยนี้จึงได้อาศัยความแตกต่างของคุณลักษณะของภาพเส้นด้าที่ปรากฏบนตัวของกุ้ง เพื่อทำการตรวจสอบการกำจัดเส้นด้าที่ไม่เหมาะสมในกุ้งแช่แข็งตามเกณฑ์มาตรฐาน

### 3.2 การเตรียมตัวอย่าง และการออกแบบระบบดึงภาพ

การเตรียมตัวอย่างกุ้งเป็นขั้นตอนสำคัญที่ช่วยให้การวิเคราะห์ด้วยเทคนิคประมวลผลภาพมีประสิทธิภาพ โดยต้องเตรียมตัวอย่างกุ้งที่มีเส้นด้าขนาดความยาวต่างๆ เพื่อให้ครอบคลุมลักษณะที่ตรวจสอบ ตัวอย่างกุ้งเหล่านี้ถูกเก็บภาพจากการทดลองด้วยระบบดึงภาพในห้องวิจัย ระบบดึงภาพได้ถูกออกแบบขึ้นเพื่อเป็นการสร้างสภาพแวดล้อมที่เหมาะสมในการเก็บข้อมูลภาพ และเพื่อจัดสภาวะให้ง่ายต่อการประมวลผล ลดสิ่งรบกวนที่จะปรากฏบนภาพถ่าย ซึ่งมีรายละเอียดดังนี้

### 3.2.1 ตัวอย่างกุ้ง (Samples)

การศึกษาวิจัยนี้ใช้ตัวอย่างกุ้งขาว (*Litopenaeus vanamei*) ซึ่งเป็นสายพันธุ์กุ้งที่มีสัดส่วนของปริมาณการใช้เป็นวัตถุดิบในผลิตภัณฑ์กุ้งสดแช่แข็งมากที่สุด โดยใช้ตัวอย่างกุ้งที่มีการผ่าหลังกำจัดเส้นดำออกจำนวน 200 ตัวอย่าง ประกอบด้วยกุ้งที่ผ่านเกณฑ์มาตรฐาน (passed) การกำจัดเส้นดำจำนวน 31 ตัวอย่างและกุ้งไม่ผ่านเกณฑ์มาตรฐาน (not passed) การกำจัดเส้นดำจำนวน 169 ตัวอย่าง ซึ่งมาตรฐานการนำเข้าผลิตภัณฑ์กุ้งของสหรัฐอเมริกาได้ระบุจำนวนตัวอย่างสำหรับการแบ่งเกรดว่าสำหรับกุ้งที่มีขนาดตั้งแต่ 70 ตัวต่อปอนด์ขึ้นไปจะมีการสุ่มตัวแทนเพื่อตรวจสอบเกรดในหนึ่งล็อตสินค้า 2 ปอนด์หรือ 0.91 กิโลกรัม และในระหว่างรอการบันทึกภาพถ่ายกุ้งจะถูกเก็บรักษาไว้ในตู้แช่เยือกแข็งอุณหภูมิ-18 องศาเซลเซียส โดยแบ่งข้อมูลภาพทดสอบออกเป็น 2 กลุ่ม ได้แก่ ข้อมูลภาพที่ใช้ในการสอน (training) และข้อมูลภาพที่ใช้ในการทดสอบ (testing)

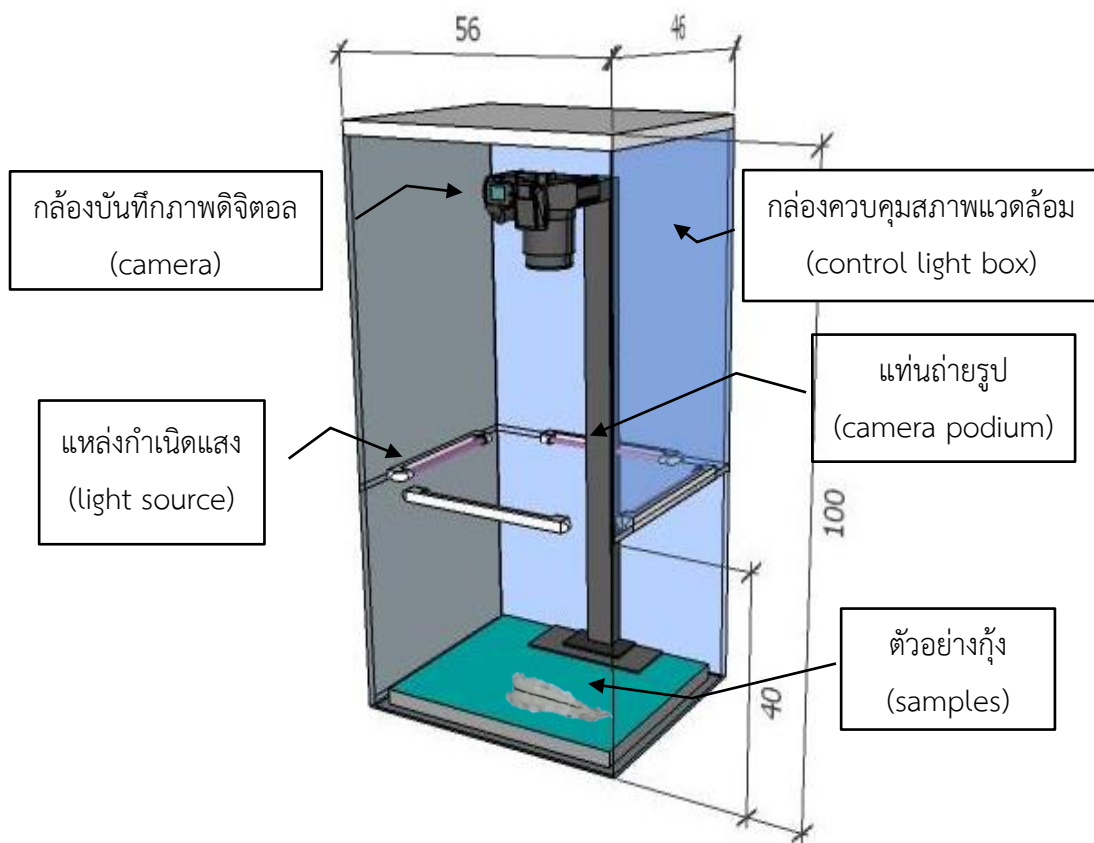


รูปที่ 3.2 กุ้งที่มีลักษณะเส้นดำตามเกณฑ์มาตรฐานการกำจัดเส้นดำ  
(ก) ไม่ผ่านเกณฑ์มาตรฐาน (ข) ผ่านเกณฑ์มาตรฐาน

จากรูปที่ 3.2 แสดงถึงรูปกุ้งที่แบ่งลักษณะของเส้นดำตามความยาวของข้อกุ้งต่างๆ เกณฑ์มาตรฐานได้ระบุความยาวของเส้นดำเปรียบเทียบกับขนาดความยาวของข้อกุ้ง โดยกลุ่มของกุ้งที่ไม่ผ่านเกณฑ์มาตรฐานจะมีความยาวรวมทั้งหมดของเส้นดำตั้งแต่เต็มลำตัวจนถึงหนึ่งข้อลำตัวดังกลุ่ม (ก) ในรูปที่ 3.2 และกลุ่มของกุ้งที่ผ่านเกณฑ์มาตรฐานจะมีความยาวรวมทั้งหมดของเส้นดำไม่เกินหนึ่งข้อลำตัวดังกลุ่ม (ข) ในรูปที่ 3.2

### 3.2.2 วัสดุอุปกรณ์ และการจัดสภาพแวดล้อม

ระบบดึงภาพสำหรับการเก็บข้อมูลภาพกึ่งที่มีการกำจัดเส้นดำในลักษณะต่างๆ ประกอบไปด้วยส่วนของวัสดุอุปกรณ์สำหรับการดึงข้อมูลภาพ ซึ่งมีการจัดสภาพแวดล้อมสำหรับการเก็บภาพโดยมีการควบคุมสภาวะต่างๆ เช่น ค่าของกล้อง แสงสว่าง พื้นหลัง เป็นต้น เพื่อให้ได้ภาพถ่ายของตัวอย่างกึ่งที่เหมาะสมต่อการนำไปวิเคราะห์ และสะดวกต่อการประมวลผลดังแสดงในรูปที่ 3.3 โดยมีส่วนประกอบและรายละเอียดต่างๆ ดังนี้



รูปที่ 3.3 แบบจำลองระบบดึงข้อมูลภาพ (หน่วยเป็นเซนติเมตร)

1) กล้องบันทึกภาพดิจิทัล (camera) ยี่ห้อ Nikon รุ่น D5100 แบบสะท้อนเลนส์เดี่ยวบันทึกภาพด้วยความละเอียด 16.2 ล้านพิกเซล เลนส์ EF-S ความยาวโฟกัส 18-55 mm. รูรับแสง F3.6-5.6 (รูปที่ 3.4) ปรับตั้งโหมดการบันทึกภาพในระบบสี sRGB มีการตั้งค่าต่างๆ ของกล้อง ดังตารางที่ 3.1 และเลนส์ทำมุม 0 องศากับตัวอย่าง โดยกล้องจะติดตั้งในแนวตั้งเหนือพื้นหลังที่เป็นแผ่นโฟมสีฟ้า ซึ่งแผ่นโฟมจะสามารถป้องกันการสะท้อนของแสงและเมื่อวางตัวอย่างที่มีความชื้นจะไม่ทำให้พื้นหลังมีสีเปลี่ยนไป



รูปที่ 3.4 กล้องบันทึกภาพดิจิทัล

ตารางที่ 3.1 การกำหนดค่าเริ่มต้นของกล้องบันทึกภาพดิจิทัล

รายละเอียด	สถานะ
ระยะทางโฟกัสของเลนส์	38 มิลลิเมตร
ระยะโฟกัส	โฟกัสอัตโนมัติ
ขนาดภาพ	2,464 × 1,632 pixels
แฟลช	ไม่ใช้
สมดุลแสงสีขาว	Daylight
ความเร็วชัตเตอร์	1/50 วินาที
ขนาดรูรับแสง	1/3"
วิธีการถ่ายภาพ	ถ่ายภาพแบบช่วงเวลา

2) แทนถ่ายรูป (camera podium) ยี่ห้อ KAISER RSI สำหรับติดตั้งกล้องบันทึกภาพ มีขนาดความยาว 45 เซนติเมตร ความกว้าง 50 เซนติเมตร และสามารถปรับระยะความสูงของกล้องตามลักษณะการใช้งานได้ตั้งแต่ 20 ถึง 90 เซนติเมตร ดังแสดงในรูปที่ 3.5



รูปที่ 3.5 แท่นถ่ายรูป ยี่ห้อ KAISER RSI

3) กล่องควบคุมสภาพสิ่งแวดล้อม (control light box) ได้ถูกสร้างขึ้นเพื่อควบคุมสภาวะแสง และสภาพแวดล้อมต่างๆในการถ่ายภาพ เพื่อให้ถ่ายต่อวิเคราะห์ข้อมูลภาพ เนื่องจากสามารถลดและป้องกันแสงรบกวนจากภายนอก อีกทั้งสัญญาณรบกวนของภาพได้อีกด้วย กล่องควบคุมสภาพสิ่งแวดล้อมที่ใช้ในงานวิจัยนี้เป็นกล่องที่ทำจากกระดานแข็ง (hard board) สร้างเป็นสี่เหลี่ยมปิดทุกด้าน ขนาดกว้าง 46 เซนติเมตร ยาว 56 เซนติเมตร และสูง 100 เซนติเมตรดังรูปที่ 3.3 โดยออกแบบให้สามารถครอบคลุมขนาดของแท่นถ่ายรูป ซึ่งมีขนาดความกว้าง 40 เซนติเมตร

4) แหล่งกำเนิดแสง (light source) งานวิจัยนี้ได้ใช้หลอดไฟลูออเรสเซนต์ชนิด Illuminant D (D65) ขนาดความยาว 30 เซนติเมตรดังรูปที่ 3.6 ซึ่งเป็นแหล่งกำเนิดแสงที่มีค่าอุณหภูมิสีของแสงเท่ากับค่าข้างต้น และให้ค่าสีที่ได้ใกล้เคียงค่าสีที่รับรู้โดยมนุษย์มากที่สุด จำนวน 4 หลอด ติดตั้งตามแนวความยาวของโครงสร้างกล่องควบคุมสภาพสิ่งแวดล้อมทั้ง 4 ด้าน ซึ่งทำมุม 90 องศา กับแนวแกนของกล่อง และแสงส่องสว่างครอบคลุมไปยังวัตถุอย่างสม่ำเสมอ



รูปที่ 3.6 หลอดไฟฟลูออเรสเซนต์ชนิด Illuminant D (D65)

#### 5) โปรแกรมสำหรับพัฒนาเทคนิคการประมวลผลภาพ ประกอบด้วย

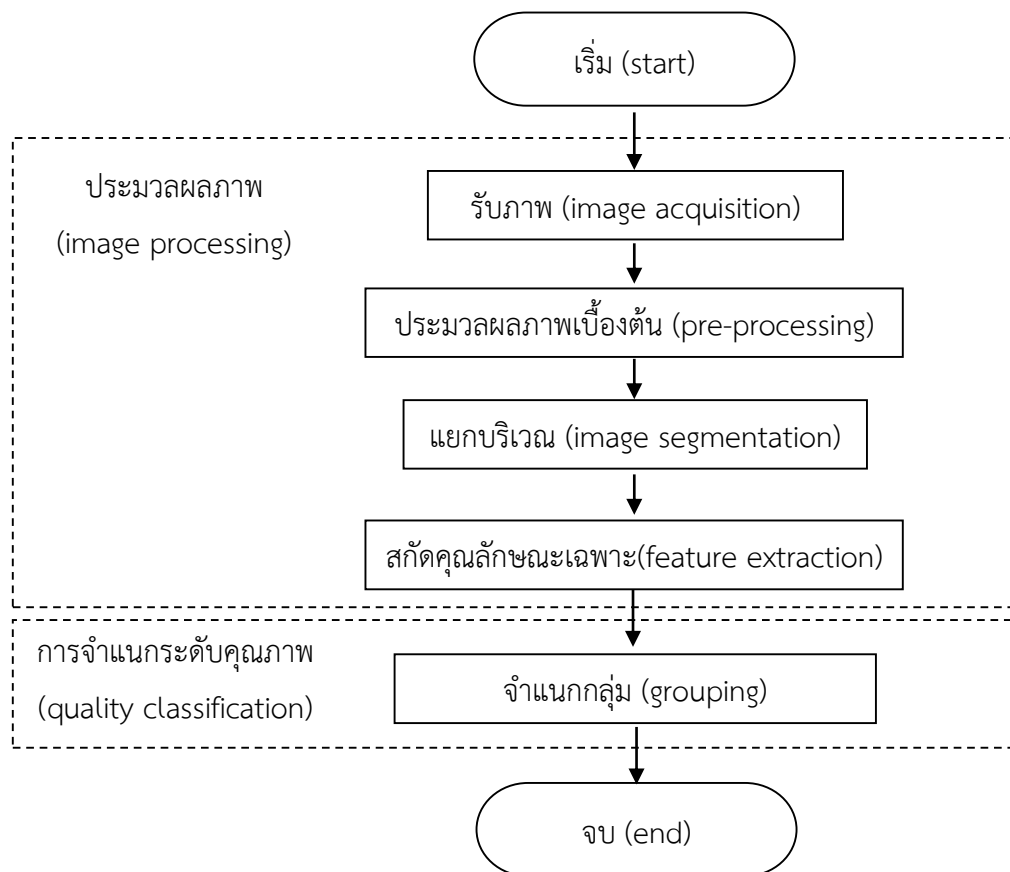
5.1) โปรแกรม ImageJ เป็นโปรแกรมฟรีแวร์ที่ใช้ในการวิเคราะห์ภาพถ่าย สำหรับงานวิจัยนี้ใช้โปรแกรม Image J เวอร์ชัน 1.51k เพื่อรับภาพ แปลงประเภทภาพเป็นภาพระดับเทา แบ่งส่วนภาพ (image segmentation) เฉพาะในบริเวณภาพที่สนใจ (Region of Interest: ROI) จากนั้นวิเคราะห์ภาพเพื่อดึงข้อมูลความยาว พื้นที่ และค่าคุณลักษณะต่างๆจากฮิสโตแกรม

5.2) โปรแกรม Weka เป็นซอฟต์แวร์สำเร็จประเภทฟรีแวร์ใช้ในเทคนิคด้านการเรียนรู้ด้วยเครื่อง (Machine Learning) และการทำเหมืองข้อมูล (Data Mining) สำหรับงานวิจัยนี้ใช้โปรแกรม Weka เวอร์ชัน 3.8.1 ดังรูปที่ 3.8 ในการเตรียมข้อมูล (Pre-Processing) ด้วยการทำนอร์มัลไลเซชัน (Normalization) เพื่อช่วยลดความซ้ำซ้อนของข้อมูลจากฐานข้อมูล โดยการกำหนดแอตทริบิว (Attribute) ให้กับแต่ละเอนทิตี (Entity) เพื่อให้สามารถควบคุมความซ้ำซ้อนของข้อมูลและหลีกเลี่ยงความผิดปกติของข้อมูล และจากนั้นนำข้อมูลมาพัฒนาการจำแนกลักษณะบ่งพร่องในกุ่มด้วยการใช้อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine หรือ SVM)

5.3) โปรแกรมวิเคราะห์ข้อมูลทางสถิติ SPSS เวอร์ชัน 22.0 เป็นโปรแกรมวิเคราะห์ และการจัดการข้อมูลโดยใช้สถิติประเภทต่างๆให้ผลลัพธ์ที่รวดเร็ว และถูกต้องแม่นยำ ซึ่งในงานวิจัยนี้โปรแกรม SPSS มาใช้ในการวิเคราะห์การจัดกลุ่ม (Discriminant Analysis) ของค่าคุณลักษณะที่ได้จากภาพถ่าย รวมทั้งใช้ในการลดขนาดของข้อมูลด้วยวิธี Principal Component Analysis (PCA)

### 3.3 การพัฒนาเทคนิคการประมวลผลภาพสำหรับการตรวจสอบข้อบกพร่องในกึ่ง

งานวิจัยนี้เป็นการพัฒนาเทคนิคประมวลผลภาพเพื่อใช้ในการตรวจสอบข้อบกพร่องของกระบวนการกำจัดเส้นดำในผลิตภัณฑ์กึ่งสแตนเลส การตรวจสอบนี้เป็นการตรวจสอบขนาดของเส้นดำว่าเป็นไปตามเกณฑ์มาตรฐานหรือไม่ โดยแบ่งกระบวนการออกเป็น 2 ส่วนดังรูปที่ 3.9 คือ ส่วนแรกเป็นการประมวลผลภาพเพื่อแยกคุณลักษณะที่สำคัญ และส่วนที่สองเป็นกระบวนการจำแนกกลุ่มโดยใช้ข้อมูลคุณลักษณะที่ได้รับจากกระบวนการแรก โดยทั้งสองส่วนมีรายละเอียดดังนี้



รูปที่ 3.7 ขั้นตอนการพัฒนาเทคนิคการประมวลผลภาพสำหรับการตรวจสอบ

### 3.3.1 การประมวลผลภาพ (image processing)

ส่วนที่หนึ่งเป็นการนำเทคนิคประมวลผลภาพมาประยุกต์ใช้แทนการมองเห็นด้วยตาของมนุษย์ เพื่อดึงข้อมูลที่ใช้สำหรับการจำแนกกลุ่มคุณภาพของภาพตัวอย่างประกอบด้วย 4 ขั้นตอน ได้แก่ การรับภาพ การประมวลผลภาพเบื้องต้น การแยกบริเวณภาพ และการสกัดคุณลักษณะเฉพาะ มีรายละเอียดดังต่อไปนี้

#### 3.3.1.1 การรับภาพ (image acquisition)

ลักษณะข้อบกพร่องของการกำจัดเส้นดำตามการเกณฑ์ตรวจสอบคุณภาพของมาตรฐานในงานวิจัยนี้ได้ถูกบันทึกด้วยระบบดึงข้อมูลภาพซึ่งมีรายละเอียดของระบบและการกำหนดค่าต่างๆของกล้องที่ใช้ในการถ่ายภาพในหัวข้อที่ 3.1 และภาพจากการถ่ายได้ถูกบันทึกจัดเก็บเป็นภาพสี RGB รูปแบบไฟล์ภาพ JPEG ในเมมโมรี่การ์ดก่อนส่งเข้าสู่คอมพิวเตอร์เพื่อดำเนินการขั้นตอนต่อไป

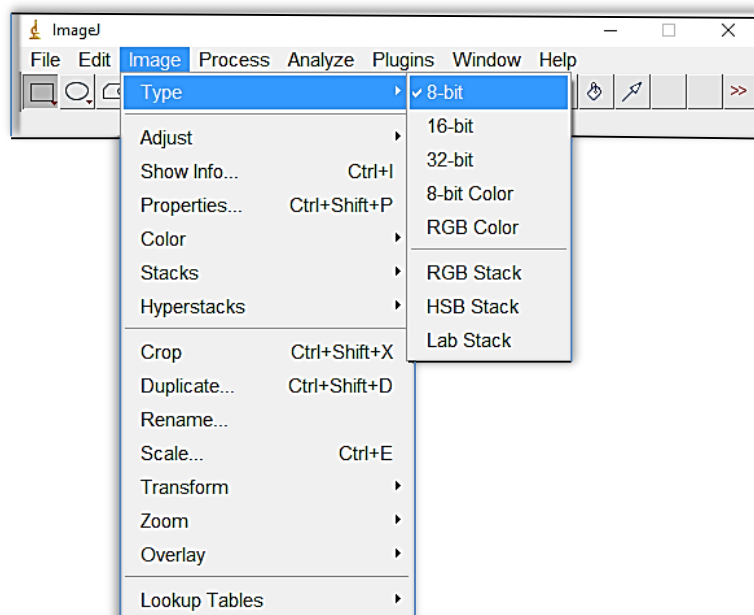
#### 3.3.1.2 การประมวลผลภาพเบื้องต้น (pre-processing)

การประมวลผลภาพเบื้องต้นเป็นขั้นตอนปรับปรุงคุณภาพของข้อมูลภาพถ่ายให้มีความเหมาะสมก่อนนำไปประมวลผลภาพจริง โดยภาพตัวอย่างกึ่งที่รับเข้ามาในตอนแรกจะเป็นภาพที่อยู่ในระบบสีแบบ RGB แต่ละพิกัดของภาพจะประกอบด้วยค่าของเซตความเข้มที่ค่าของค่า R ค่า G และค่า B จากการสำรวจเอกสารพบว่า เส้นดำหรือลำไส้ซึ่งเป็นส่วนหนึ่งของระบบย่อยอาหารของกึ่งส่วนใหญ่จะมีสีดำ แต่ความเข้มอาจแตกต่างกันขึ้นอยู่กับลักษณะอาหารที่กึ่งกิน งานวิจัยนี้ได้อาศัยความแตกต่างของคุณลักษณะสีของเส้นดำที่ปรากฏบนตัวของกึ่ง ทำให้ค่าสีอื่นๆในภาพไม่ส่งผลต่อการวิเคราะห์ ดังนั้นจึงใช้เทคนิคการแปลงภาพสีแบบ RGB เป็นภาพระดับเทา ภาพจะแสดงเฉพาะค่าความสว่าง (brightness) เพียงอย่างเดียวโดยไม่มีค่าสีของภาพ ทำให้สามารถวิเคราะห์ภาพได้ง่ายขึ้นเนื่องจากเมื่อแปลงภาพเป็นระดับเทาแล้วจะทำให้แต่ละจุดภาพของภาพเหลือเพียงค่าความสว่าง 256 ระดับ ซึ่งเป็นค่าตั้งแต่ 0 ถึง 255 ทำได้โดยแยกระดับสีแต่ละพิกเซลในภาพระบบสีแบบ RGB จากนั้นนำค่าสีดังกล่าวมาคำนวณด้วยด้วยสมการที่ 2.2 ซึ่งจะได้ค่าสีระดับเทาเพื่อนำไปแทนในจุดพิกเซลเดิม

การแปลงระบบสีของภาพโดยใช้โปรแกรม ImageJ เวอร์ชัน 1.51k ชุดคำสั่งแสดงดังรูปที่ 3.8 โดยมีขั้นตอนในการแปลงภาพสี RGB เป็นภาพระดับเทาอยู่ที่เมนู



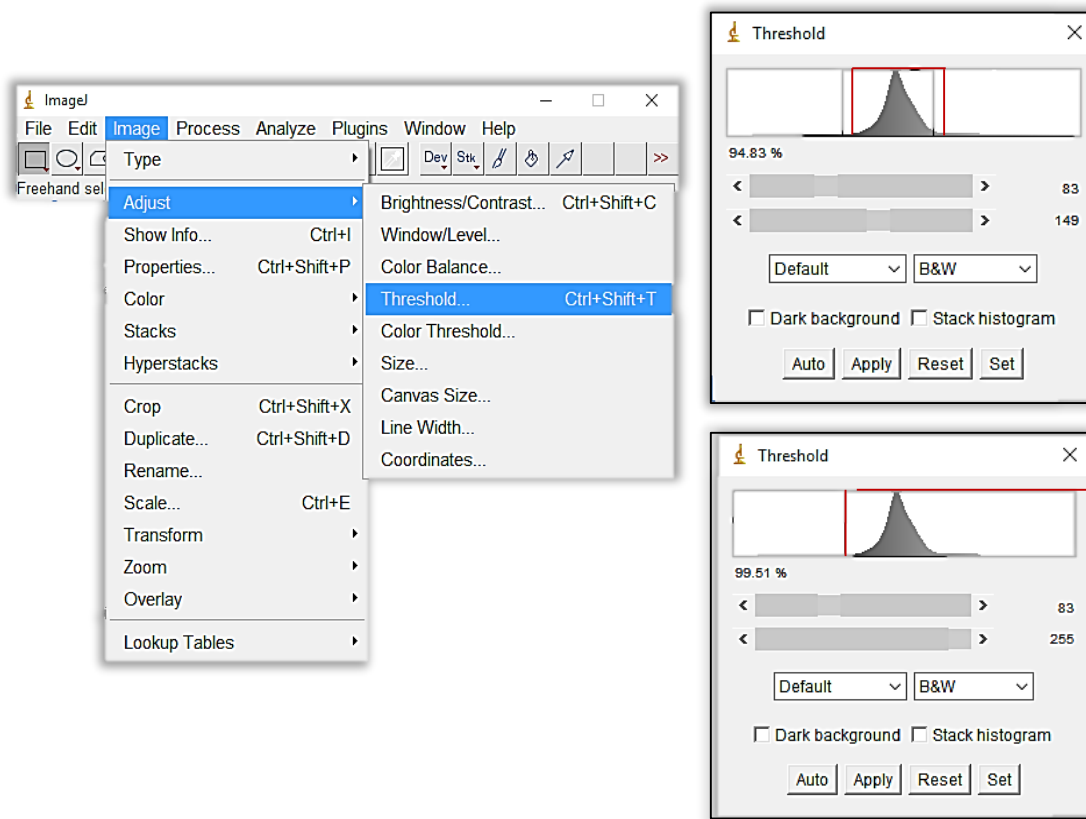
Image → Type → 8-bit ซึ่งภาพระดับเทาเรียกอีกชื่อหนึ่งว่า ภาพ 8 บิต เนื่องจากมีการแสดงค่าความเข้มสีได้  $2^8$



รูปที่ 3.8 คำสั่งในโปรแกรม ImageJ สำหรับแปลงภาพสี RGB เป็นภาพระดับเทา

### 3.3.1.3 การแยกบริเวณ (image segmentation)

การแยกบริเวณเป็นขั้นตอนการแยกส่วนของภาพเพื่อแยกวัตถุที่สนใจออกจากพื้นหลัง โดยวัตถุที่แยกออกมานี้มีส่วนสำคัญอยู่สองบริเวณ นั่นคือส่วนลำตัวและส่วนเส้นดำของกึ่ง โดยจะใช้วิธีเทคนิคการแบ่งส่วนภาพด้วยการขีดแบ่ง (thresholding) ซึ่งเป็นวิธีที่นิยมใช้ในหลายงานวิจัย ความเข้มในภาพระดับเทาที่มีค่าน้อยกว่าค่าแบ่ง (thresholding value) จะถูกแปลงเป็น 0 ซึ่งจะแสดงผลเป็นความมืด และค่าที่สูงกว่าค่าแบ่งจะถูกแปลงเป็นค่าสูงค่าหนึ่งที่จะแสดงผลเป็นความสว่าง ภาพถ่ายตัวอย่างกึ่งมีการใช้ค่าแบ่งหลายค่า เนื่องจากกึ่งเป็นวัตถุที่มีสีไม่สม่ำเสมอแต่ละส่วนจะมีช่วงกว้างของความเข้มต่างกัน ซึ่งสามารถทำการแยกบริเวณภาพด้วยค่าแบ่งหาได้จากฮิสโตแกรมของภาพแต่ละภาพ โดยชุดคำสั่งและการกำหนดค่าแบ่งทั้ง 2 ค่าดังตัวอย่างรูปที่ 3.9 ผลลัพธ์จากขั้นตอนนี้จะทำให้บริเวณตัว และเส้นดำกึ่งถูกแยกออกจากพื้นหลังโดยยังคงแสดงค่าคุณลักษณะความเข้มจริงของภาพกึ่งเพื่อใช้ในการแยกคุณลักษณะในขั้นตอนต่อไป

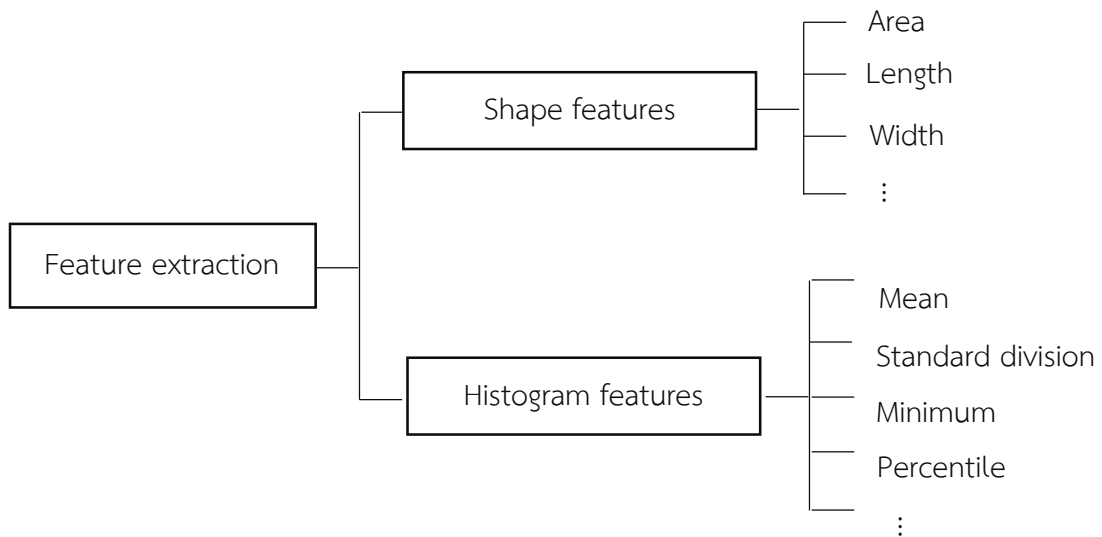


รูปที่ 3.9 ชุดคำสั่งการแบ่งส่วนภาพด้วยการขีดแบ่งและปรับค่าแบ่ง 2 ค่าจากฮิสโตแกรมของภาพ

จากรูปที่ 3.9 การแยกบริเวณภาพมีชุดคำสั่งที่เมนู Image → Adjust → Threshold และปรับค่าแบ่งสองค่าให้เหมาะสมตามฮิสโตแกรมของภาพ จากตัวอย่างค่าแบ่งส่วนตัวกึ่งใช้เป็นค่าระหว่าง 83 ถึง 149 ซึ่งช่วงที่มีพื้นที่ของค่าความเข้มสูง และส่วนเส้นดำมีค่าแบ่งระหว่าง 83 ถึง 255 เมื่อใช้คำสั่งดังกล่าวในการแบ่งส่วนภาพด้วยการขีดแบ่งเพื่อหาค่าที่เหมาะสมในแต่ละภาพจะได้รูปตัวอย่างกึ่งที่แยกบริเวณออกมา

#### 3.3.1.4 การสกัดคุณลักษณะเฉพาะ (feature extraction)

การสกัดคุณลักษณะเฉพาะของภาพเป็นกระบวนการตรวจจับหรือดึงลักษณะที่สำคัญของวัตถุออกมา ในงานวิจัยนี้มุ่งเน้นไปที่การตรวจจับเส้นดำที่เป็นขอบกพร่องในผลิตภัณฑ์กึ่ง คุณลักษณะของเส้นดำที่คาดว่าจะมีส่วนในการจำแนกกลุ่มตามเกณฑ์มาตรฐานประกอบด้วยคุณลักษณะดังแสดงในรูปที่ 3.10 โดยมีรายละเอียดและขั้นตอนในการดึงข้อมูลในแต่ละประเภทดังนี้



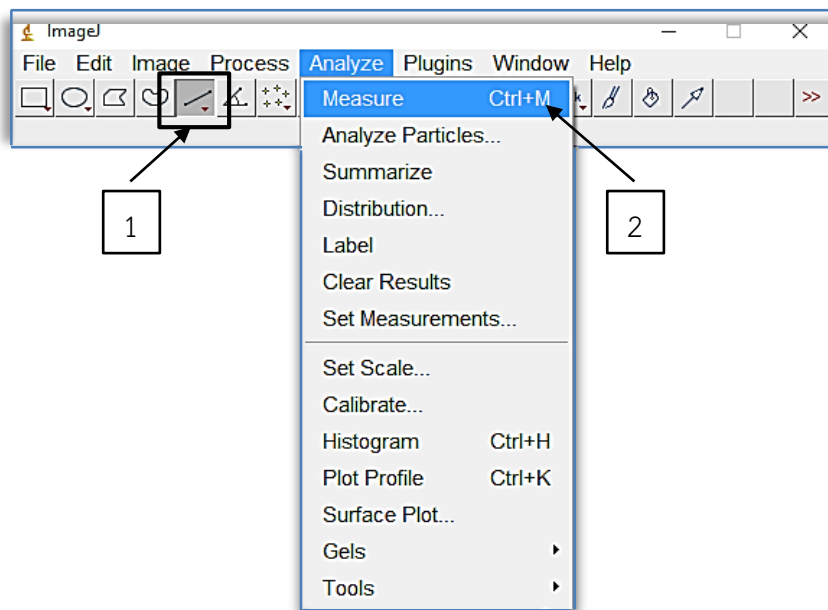
รูปที่ 3.10 คุณลักษณะเฉพาะที่สกัดได้จากภาพ

### 1) คุณลักษณะทางรูปร่าง (shape features)

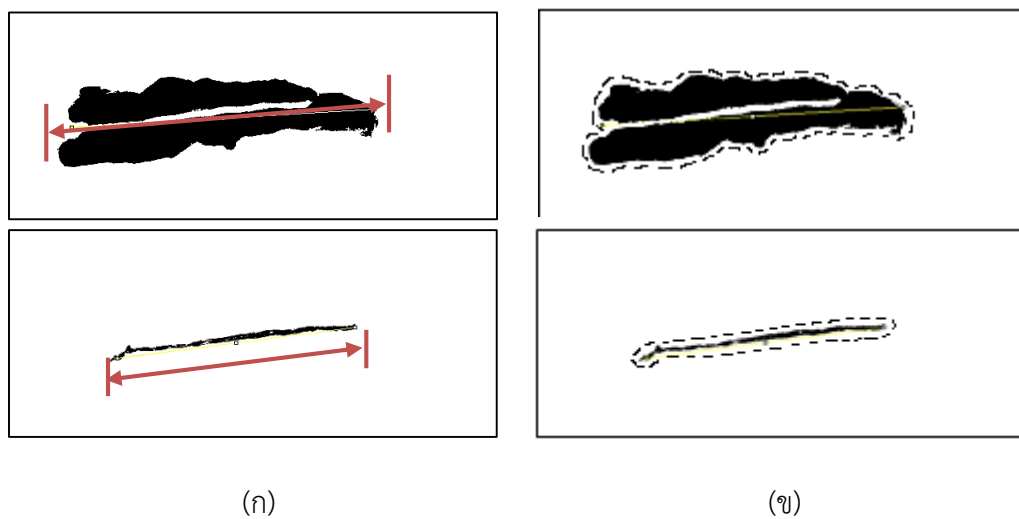
เส้นดำเป็นส่วนประกอบของกึ่งที่สามารถมองเห็นได้อย่างชัดเจนจากภาพถ่าย ดังนั้นคุณลักษณะแรกของเส้นดำที่สามารถแยกออกมาได้คือคุณลักษณะทางรูปร่าง ซึ่งสามารถแยกออกมาได้เป็น พื้นที่ และความยาว โดยในงานวิจัยนี้ใช้การเปรียบเทียบสัดส่วนระหว่างขนาดตัวกึ่งกับเส้นดำของกึ่ง ดังนี้

ก) ความยาว (length : L) สามารถวัดจากความยาวของเส้นดำที่ปรากฏ ( $L_v$ ) และความยาวของตัวกึ่งตั้งแต่ส่วนหัวถึงส่วนหาง ( $L_b$ ) โดยใช้เครื่องมือ Straight ดังลูกศรชี้หมายเลข 1 ในรูปที่ 3.11 ลากเส้นในการวัดตามความยาวของแต่ละภาพดังแสดงแสดงในรูปที่ 3.12 (ก) จากนั้นใช้ชุดคำสั่ง Analyze  $\rightarrow$  Measure เพื่อประมวลผลความยาวจากการนับจำนวนพิกเซลที่เครื่องมือได้ลากเส้นผ่าน แล้วจึงนำค่าที่ได้เทียบหาสัดส่วนด้วยสมการที่ 3.2

ข) พื้นที่ (Area) สามารถวัดพื้นที่ที่ได้จากการใช้ชุดคำสั่ง Analyze  $\rightarrow$  Measure ดังลูกศรชี้หมายเลข 2 ในรูปที่ 3.11 โดยไม่ต้องใช้เครื่องมือในการลากเส้นกำหนดขอบเขต เนื่องจากได้ทำการแบ่งส่วนที่สนใจด้วยเทคนิคการการแบ่งส่วนภาพด้วยการขีดแบ่งแล้ว ดังนั้นจะได้ค่าของพื้นที่จากการประมวลผลการรวบรวมพื้นที่ของแต่ละพิกเซลของบริเวณเส้นดำ ( $A_v$ ) และบริเวณตัวกึ่งทั้งหมด ( $A_b$ ) ซึ่งอยู่ในเส้นประดังรูปที่ 3.12 (ข) แล้วจึงนำค่าที่ได้เทียบหาสัดส่วนด้วยสมการที่ 3.3



รูปที่ 3.11 ชุดคำสั่งในการแยกคุณลักษณะทางรูปร่าง



รูปที่ 3.12 การวัดคุณลักษณะทางรูปร่าง  
(ก) ความยาวลำตัว และเส้นดำ (ข) พื้นที่ลำตัว และเส้นดำ

สมการสัดส่วนความยาว :

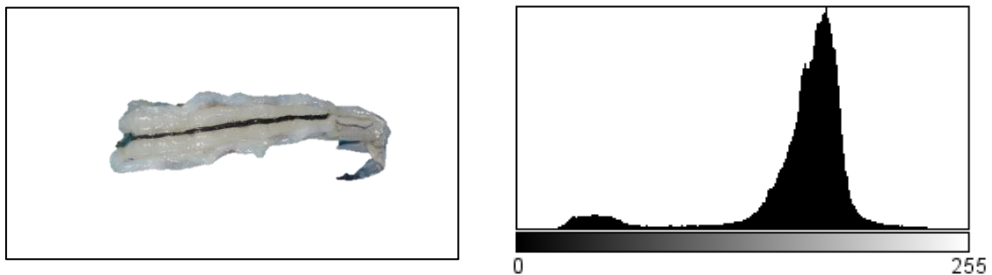
$$L = \frac{L_v}{L_b} \quad (3.2)$$

สมการสัดส่วนพื้นที่ :

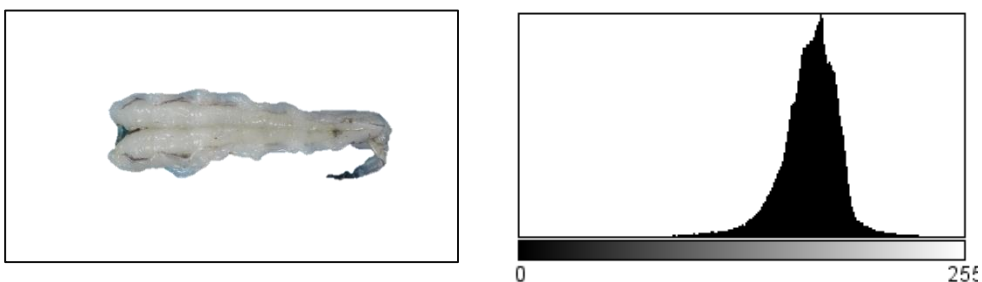
$$A = \frac{A_v}{A_b} \quad (3.3)$$

## 2) คุณลักษณะจากฮิสโตแกรม (histogram features)

ฮิสโตแกรมเป็นแผนภูมิแสดงความถี่หรือจำนวนของพิกเซลที่กระจายอยู่ ณ ช่วงค่าความสว่างของภาพตั้งแต่ 0-255 (256 ระดับ) เมื่อนำภาพกึ่งที่ผ่านกระบวนการแยกบริเวณมาวิเคราะห์ด้วยฮิสโตแกรม พบว่าฮิสโตแกรมของภาพกึ่งที่มีลักษณะของเส้นดำผ่านเกณฑ์มาตรฐานมีความแตกต่างเมื่อเทียบกับฮิสโตแกรมของภาพกึ่งในกลุ่มไม่ผ่านเกณฑ์มาตรฐานดังแสดงในรูปที่ 3.13 และ 3.14 โดยคุณลักษณะที่นิยมใช้ประกอบด้วยค่าเฉลี่ย (mean) ค่าความสว่างสูงสุด (maximum) ค่าความสว่างต่ำสุด (minimum) ฐานนิยม (mode) ค่าเบี่ยงเบนมาตรฐาน (standard division) ความเบ้ (skewness) ความโด่ง (kurtosis) และจำนวนของความถี่หรือพิกเซลสะสม ณ เปอร์เซ็นไทล์ที่ตำแหน่งต่างๆ (percentile) งานวิจัยนี้เป็นการตรวจจับเส้นดำที่มีขนาดเกินเกณฑ์ข้อกำหนด จากการทดลองเบื้องต้น (preliminary experiment) กับกลุ่มตัวอย่างทดสอบจำนวน 100 ตัวอย่าง พบว่าคุณลักษณะที่ส่งผลต่อการจัดกลุ่มข้อบกพร่อง ได้แก่ ค่าเฉลี่ยของค่าความสว่าง ค่าความสว่างที่ต่ำสุด ค่าเบี่ยงเบนมาตรฐาน และค่าของเปอร์เซ็นไทล์ที่ 25 , 50 และ 75



รูปที่ 3.13 กุ้งและฮิสโตแกรมของกุ้งที่ไม่ผ่านเกณฑ์มาตรฐาน



รูปที่ 3.14 กุ้งและฮิสโตแกรมของกุ้งที่ผ่านเกณฑ์มาตรฐาน

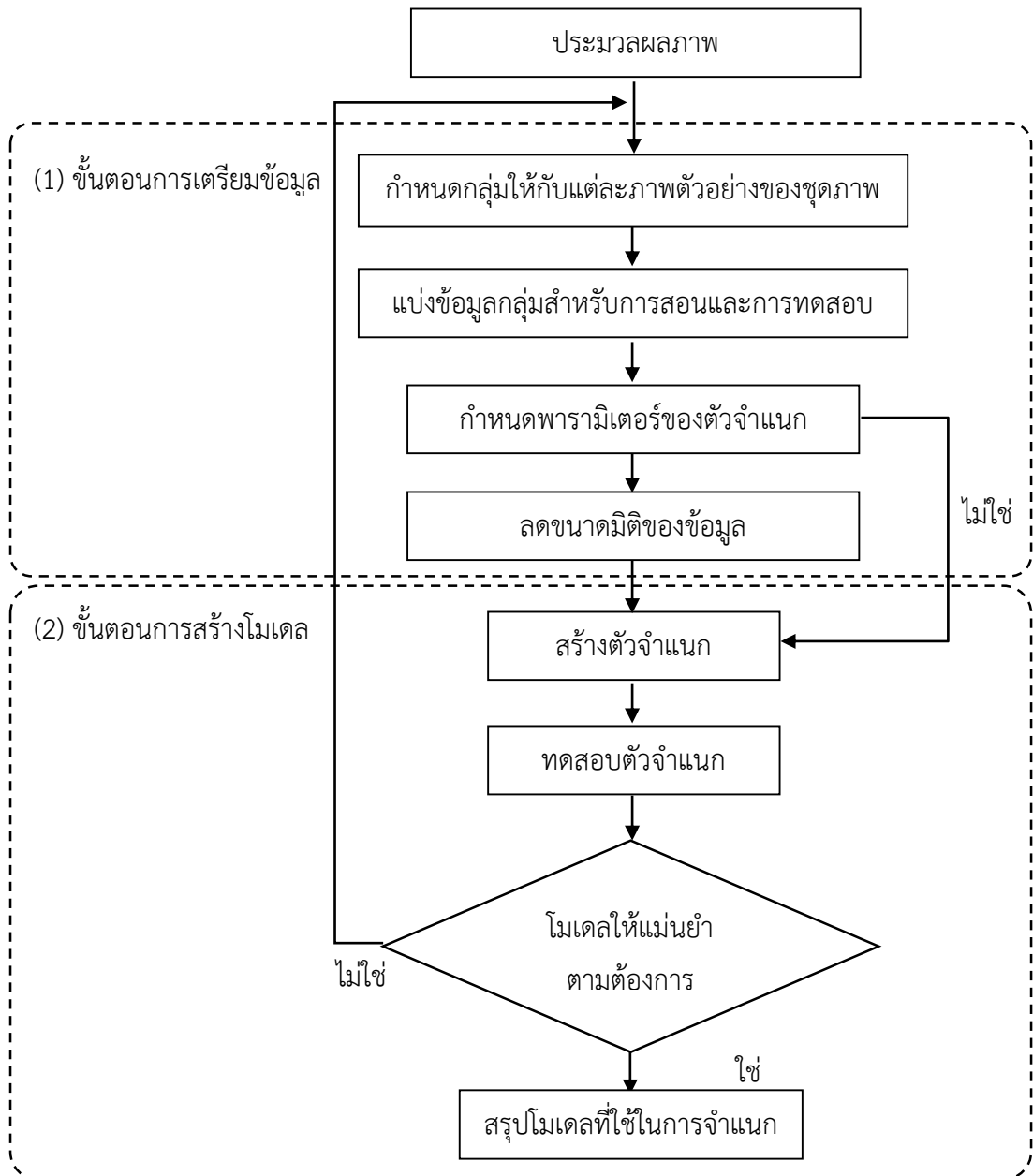
เมื่อดึงคุณลักษณะเฉพาะของภาพกึ่งออกมาแล้ว จะใช้คุณลักษณะดังกล่าวเป็นตัวแปรอิสระหรือพารามิเตอร์ในการจำแนก ดังนั้นเพื่อความเข้าใจและความสะดวกในการวิเคราะห์จึงกำหนดชื่อตัวแปรแทนคุณลักษณะดังตารางที่ 3.1 ต่อไป

ตารางที่ 3.2 ชื่อตัวแปรและคุณลักษณะ

ลำดับ	คุณลักษณะ	ชื่อตัวแปร
1	สัดส่วนพื้นที่	พื้นที่ หรือ Area
2	สัดส่วนความยาว	ความยาว หรือ Length
3	ค่าเฉลี่ย	ค่าเฉลี่ย หรือ Mean
4	ส่วนเบี่ยงเบนมาตรฐาน	ส่วนเบี่ยงเบนมาตรฐาน หรือ Std
5	ค่าความสว่างต่ำสุด	ค่าต่ำสุด หรือ Min
6	จำนวนพิกเซล ณ เเปอร์เซ็นต์ไทล์ที่ 25	เปอร์เซ็นต์ไทล์ที่ 25 หรือ Per 25
7	จำนวนพิกเซล ณ เเปอร์เซ็นต์ไทล์ที่ 50	เปอร์เซ็นต์ไทล์ที่ 50 หรือ Per 50
8	จำนวนพิกเซล ณ เเปอร์เซ็นต์ไทล์ที่ 75	เปอร์เซ็นต์ไทล์ที่ 75 หรือ Per 75

### 3.3.2 การจำแนกระดับคุณภาพ (quality classification)

การจำแนกระดับคุณภาพเป็นกระบวนการที่นำข้อมูลที่ได้จากการประมวลผลภาพมาวิเคราะห์เพื่อจำแนกกลุ่มตามเกณฑ์มาตรฐาน โดยในงานวิจัยนี้ประกอบด้วย 2 ขั้นตอน ดังรูปที่ 3.15 ซึ่งประกอบด้วย ขั้นตอนแรกเป็นการเตรียมข้อมูล (preparing phase) และขั้นตอนที่สองเป็นการสร้างโมเดลเพื่อนำไปใช้ในการจำแนกระดับคุณภาพ (modeling phase) มีรายละเอียดแต่ละขั้นตอนดังนี้



รูปที่ 3.15 กระบวนการจำแนกกลุ่มข้อบกพร่อง

### 3.3.2.1 ขั้นตอนการเตรียมข้อมูล (preparing phase)

ในขั้นตอนนี้เป็นกระบวนการก่อนนำข้อมูลมาวิเคราะห์การจำแนกระดับคุณภาพ เพื่อเป็นการเตรียมข้อมูลให้พร้อม และเหมาะสมสำหรับการใช้กับโมเดลจำแนกต่าง ๆ โดยใช้ข้อมูลภาพของกลุ่มกึ่ง ซึ่งได้จากการเก็บภาพตัวอย่างภาพกึ่งจำนวน 200 ภาพ มีรายละเอียดขั้นตอนดังนี้

1) ประเมินคุณภาพของกิ่งในแต่ละภาพตัวอย่างของชุดภาพสำหรับการสอนและการทดสอบ โดยกลุ่มดังกล่าวเป็นระดับคุณภาพของข้อบกพร่องที่ระบุตามมาตรฐานแบ่งเป็น 2 กลุ่ม ได้แก่ กลุ่มผ่านเกณฑ์มาตรฐาน และกลุ่มไม่ผ่านเกณฑ์มาตรฐาน

2) แบ่งกลุ่มข้อมูลในการสอนและทดสอบโมเดล โดยใช้กลุ่มข้อมูลสำหรับการสอนเพื่อสร้างโมเดล และใช้กลุ่มข้อมูลสำหรับการทดสอบสำหรับพิจารณาการจำแนกคุณภาพกิ่งว่ามีความถูกต้องในการจำแนกด้วยโมเดลที่สร้างขึ้นจากขั้นตอนการสอนหรือไม่ ซึ่งในการวิจัยนี้ใช้การแบ่งข้อมูลเพื่อทำการสอนและทดสอบ 3 วิธี มีรายละเอียดดังนี้

- วิธี self-consistency test ใช้วิธีการแบ่งชุดข้อมูลนี้จะนำมาใช้ในการจำแนกระดับคุณภาพด้วยวิธี discriminant analysis และซัพพอร์ตเวกเตอร์แมชชีนโดยชุดข้อมูลที่ใช้ในการสร้างโมเดลจำแนก และชุดข้อมูลที่ใช้ในการทดสอบประสิทธิภาพของโมเดลเป็นข้อมูลชุดเดียวกัน แต่การแบ่งข้อมูลด้วยวิธีนี้จะให้ผลการวัดประสิทธิภาพที่มีค่าสูงมาก เนื่องจากเป็นข้อมูลชุดเดิมที่โมเดลได้รับการสอนมาแล้ว ซึ่งวิธีการนี้เหมาะสำหรับใช้ในการทดสอบประสิทธิภาพเพื่อดูแนวโน้มของโมเดลที่สร้างขึ้น

- วิธี split test วิธีนี้เป็นการแบ่งข้อมูลด้วยการสุ่มออกเป็น 2 ส่วน โดยข้อมูลส่วนที่หนึ่งใช้ในการสร้างโมเดลจำแนกจำนวน 140 ตัวอย่าง คิดเป็น 70% ของภาพตัวอย่างทั้งหมด และข้อมูลส่วนที่สองใช้ในการทดสอบประสิทธิภาพของโมเดล จำนวน 60 ตัวอย่าง คิดเป็น 30% ของภาพตัวอย่างทั้งหมด โดยทำการสุ่มข้อมูลเพียงหนึ่งครั้ง ทำให้การสร้างโมเดลใช้เวลาน้อยและสะดวกต่อการแบ่งชุดข้อมูล แต่วิธีการนี้ชุดกลุ่มทดสอบมีลักษณะคล้ายหรือแตกต่างกับชุดกลุ่มข้อมูลที่ใช้สร้างโมเดลมากอาจทำให้ผลการวัดประสิทธิภาพได้ออกมาผิดพลาดได้ ซึ่งวิธีการแบ่งชุดข้อมูลนี้จะนำมาใช้ในการจำแนกระดับคุณภาพด้วยวิธี discriminant analysis และวิธีซัพพอร์ตเวกเตอร์แมชชีน

- วิธี leave-one-out cross-validation test วิธีนี้เป็นวิธีแบ่งส่วนชุดข้อมูลเพื่อทดสอบประสิทธิภาพของโมเดลที่มีความน่าเชื่อถือ การวัดประสิทธิภาพด้วยวิธีนี้จะแบ่งข้อมูลให้มีค่าเท่ากับจำนวนข้อมูลทั้งหมด คือ ทำการแบ่งข้อมูลออกเป็น 2 ส่วน โดยถ้ากลุ่มตัวอย่างมีจำนวน  $N$  ตัวอย่าง ข้อมูลส่วนแรกจะมีจำนวน  $N-1$  ตัวอย่าง และข้อมูลส่วนที่สองมีจำนวน 1 ตัวอย่าง จะใช้เป็นตัวทดสอบประสิทธิภาพของโมเดลทั้งหมด  $N$  ครั้ง ใช้รูปแบบนี้แบ่งชุดข้อมูลจะนำมาใช้ในการจำแนกระดับคุณภาพด้วยวิธี Discriminant Analysis และวิธีซัพพอร์ตเวกเตอร์แมชชีนโดยแสดงดังรูปที่ 3.16





รูปที่ 3.16 วิธี leave-one-out cross-validation test

3) กำหนดพารามิเตอร์ของตัวจำแนก ซึ่งเป็นข้อมูลภาพกึ่งที่ได้จากกระบวนการประมวลผลภาพมาสกัดให้เหลือเพียงชุดของคุณลักษณะที่สำคัญและน่าสนใจเท่านั้น เพื่อนำไปใช้สำหรับสอนและจำแนกกลุ่ม โดยพิจารณาจากการนำคุณลักษณะที่ผ่านการทดสอบการแจกแจงปกติแล้วมาทดสอบเงื่อนไขความเท่ากันของความแปรปรวนด้วย F-test ก่อนทำการทดสอบทางสถิติ T-test จากกลุ่มตัวอย่างประชากร 2 กลุ่ม โดยนำค่าเฉลี่ยของแต่ละคุณลักษณะมาทำการทดสอบความแตกต่างของกลุ่มกึ่งผ่านและไม่ผ่านเกณฑ์มาตรฐาน ซึ่งตั้งสมมติฐานได้ดังตารางที่ 3.2 ที่ระดับความเชื่อมั่น 95% ( $\alpha=0.05$ ) ดังนี้

ทดสอบ F-test

$H_0$  : ความแปรปรวนของคุณลักษณะแต่ละประเภทไม่แตกต่างกัน

$H_1$  : ความแปรปรวนของคุณลักษณะแต่ละประเภทแตกต่างกัน

ทดสอบ T-test

$H_0$  : ค่าเฉลี่ยของคุณลักษณะแต่ละประเภทไม่แตกต่างกัน

$H_1$  : ค่าเฉลี่ยของคุณลักษณะแต่ละประเภทแตกต่างกัน

ตารางที่ 3.3 สมมติฐานของสถิติ F-test และ T-test

ลำดับ	คุณลักษณะ	F-test		T-test	
		H <sub>0</sub>	H <sub>1</sub>	H <sub>0</sub>	H <sub>1</sub>
1	พื้นที่	$\sigma_{11} = \sigma_{21}$	$\sigma_{11} \neq \sigma_{21}$	$\mu_{11} = \mu_{21}$	$\mu_{11} \neq \mu_{21}$
2	ความยาว	$\sigma_{12} = \sigma_{22}$	$\sigma_{12} \neq \sigma_{22}$	$\mu_{12} = \mu_{22}$	$\mu_{12} \neq \mu_{22}$
3	ค่าเฉลี่ย	$\sigma_{13} = \sigma_{23}$	$\sigma_{13} \neq \sigma_{23}$	$\mu_{13} = \mu_{23}$	$\mu_{13} \neq \mu_{23}$
4	ส่วนเบี่ยงเบนมาตรฐาน	$\sigma_{14} = \sigma_{24}$	$\sigma_{14} \neq \sigma_{24}$	$\mu_{14} = \mu_{24}$	$\mu_{14} \neq \mu_{24}$
5	ค่าต่ำสุด	$\sigma_{15} = \sigma_{25}$	$\sigma_{15} \neq \sigma_{25}$	$\mu_{15} = \mu_{25}$	$\mu_{15} \neq \mu_{25}$
6	เปอร์เซ็นต์ไทล์ที่ 25	$\sigma_{16} = \sigma_{26}$	$\sigma_{16} \neq \sigma_{26}$	$\mu_{16} = \mu_{26}$	$\mu_{16} \neq \mu_{26}$
7	เปอร์เซ็นต์ไทล์ที่ 50	$\sigma_{17} = \sigma_{27}$	$\sigma_{17} \neq \sigma_{27}$	$\mu_{17} = \mu_{27}$	$\mu_{17} \neq \mu_{27}$
8	เปอร์เซ็นต์ไทล์ที่ 75	$\sigma_{18} = \sigma_{28}$	$\sigma_{18} \neq \sigma_{28}$	$\mu_{18} = \mu_{28}$	$\mu_{18} \neq \mu_{28}$

จาก  $\mu_{ab}$  แทนความแปรปรวนของคุณลักษณะ

และ  $\mu_{ab}$  แทนค่าเฉลี่ยของคุณลักษณะ

โดยที่ a แทน กลุ่มคุณภาพกึ่ง (ผ่านมาตรฐาน แทน “1” ไม่ผ่านมาตรฐาน แทน “2”)

b แทน ลำดับของคุณลักษณะ

4) นำเทคนิคการลดขนาดข้อมูล (data reduction) มาใช้เพื่อจัดเตรียมข้อมูล ให้ข้อมูลมีขนาดลดลง โดยยังคงลักษณะเด่นและสูญเสียความถูกต้องของผลลัพธ์น้อยที่สุด ในงานวิจัยนี้ใช้วิธีลดขนาดมิติของข้อมูล (dimensionality reduction) ด้วยเทคนิค Principal Component Analysis (PCA) ของตัวอย่างภาพกึ่งจำนวน 200 ภาพ ซึ่งเป็นวิธีที่นิยมใช้ในการลดมิติของเวกเตอร์ โดยการฉายเวกเตอร์ไปบนแกนองค์ประกอบหลัก (principal component) ซึ่งมีความแตกต่างตามค่าความแปรปรวนของแต่ละแกน ผลลัพธ์ของการวิเคราะห์องค์ประกอบหลักจะได้ข้อมูลชุดใหม่ในรูปขององค์ประกอบที่เกิดจากการรวมตัวเชิงเส้น มีขั้นตอนวิเคราะห์ดังนี้

- ทดสอบข้อสมมติเบื้องต้นของการวิเคราะห์ปัจจัย (assumption)
- สร้างเมทริกซ์สัมประสิทธิ์สหสัมพันธ์ของตัวแปรทุกคู่ (correlation

matrix)

- การสกัดปัจจัย (factor extraction) ด้วยเทคนิค PCA

- หาจำนวนปัจจัยร่วมที่เหมาะสม (number of common factors)
- ให้ความหมายปัจจัยร่วม (factor meaning)
- หมุนแกนปัจจัยร่วม (factor rotation) แบบตั้งฉากกัน (orthogonal rotation) ด้วยวิธี Varimax
- คำนวณค่าคะแนนปัจจัย (factor score) โดยสมการที่ใช้ในการประมาณค่าปัจจัยหรือองค์ประกอบหลักดังสมการที่ 2.10

### 3.3.2.2 ขั้นตอนการสร้างโมเดล (modeling phase)

ขั้นตอนการสร้างโมเดลเป็นขั้นตอนการนำพารามิเตอร์ที่เป็นผลลัพธ์จากขั้นตอนการกำหนดพารามิเตอร์ และใช้เทคนิคการลดขนาดข้อมูลมาใช้ในการสร้างโมเดล โดยโมเดลที่ได้จากขั้นตอนนี้จะเป็นตัวจำแนก (classifier) ซึ่งในงานวิจัยนี้จะเปรียบเทียบผลความแม่นยำในการจำแนกระดับคุณภาพระหว่างการใช้เทคนิค 2 เทคนิค โดยมีรายละเอียดดังนี้

#### 1) การจำแนกกลุ่มด้วยเทคนิค Discriminant Analysis

วิธีนี้เป็นวิธีการทางสถิติวิธีหนึ่งที่ใช้วิเคราะห์เพื่อจำแนกกลุ่ม โดยใช้กลุ่มข้อมูลชุดสำหรับการสอนในการสร้างโมเดล มีขั้นตอนการวิเคราะห์การจำแนกกลุ่มดังนี้

##### ก) ขั้นตอนการสอน (training phase)

โดยขั้นตอนนี้ต้องตรวจสอบเงื่อนไขเบื้องต้นก่อนนำข้อมูลไปวิเคราะห์ประกอบด้วย 2 ประการดังนี้

- พารามิเตอร์มีการแจกแจงปกติหลายตัวแปร (Normality of Independent Variables) โดยตรวจสอบด้วยการทดสอบการแจกแจงแบบปกติของพารามิเตอร์แต่ละตัว ซึ่งถ้าพารามิเตอร์แต่ละตัวมีการแจกแจงแบบปกติแล้ว ก็มีความน่าจะเป็นสูงที่จะมีการแจกแจงแบบปกติหลายตัวแปรด้วย

- เมตริกซ์ความแปรปรวนร่วมของพารามิเตอร์ในแต่ละกลุ่มต้องเท่ากัน (Equal Dispersion Matrices) ตรวจสอบได้ด้วยสถิติทดสอบ Box's M

หลังจากตรวจสอบเงื่อนไขเบื้องต้นแล้ว จากนั้นทำการเตรียมข้อมูลที่จะนำมาใช้ในการจำแนกกลุ่ม โดยมีรายละเอียดดังต่อไปนี้

- แบ่งกลุ่มตัวอย่างออกเป็นกลุ่มอย่างน้อย 2 กลุ่ม โดยในงานวิจัยนี้แบ่งการจำแนกกลุ่มตามลักษณะข้อบกพร่องออกเป็น 2 กลุ่ม คือ กลุ่มที่มีการกำจัดเส้นดำไม่ผ่าน

เกณฑ์มาตรฐาน แทนด้วยกลุ่มที่ 1 และกลุ่มที่มีการกำจัดเส้นดำผ่านเกณฑ์มาตรฐาน แทนด้วยกลุ่มที่ 2

- กำหนดพารามิเตอร์เป็นคุณลักษณะที่ได้จากการแยกลักษณะของภาพ ซึ่งจากขั้นตอนดังกล่าวทำให้มีพารามิเตอร์ทั้งหมด 8 ตัวประกอบด้วย สัดส่วนของความยาว สัดส่วนของพื้นที่ ค่าเฉลี่ยของความสว่าง ค่าความสว่างที่ต่ำสุด ค่าเบี่ยงเบนมาตรฐาน และค่าของเปอร์เซ็นต์ไทล์ที่ 25 , 50 และ 75 เป็นพารามิเตอร์ที่คาดว่าจะทำให้เกิดความแตกต่างระหว่างกลุ่ม

- กำหนดค่าตัวแปรตาม ซึ่งเป็นตัวแปรที่จะถูกจำแนกแบ่งเป็น 2 กลุ่ม คือ กลุ่มที่ 1 กลุ่มที่มีการกำจัดเส้นดำไม่ผ่านเกณฑ์มาตรฐาน กำหนดให้ตัวแปรตามคือ “1” และกลุ่มที่ 2 กลุ่มที่มีการกำจัดเส้นดำผ่านเกณฑ์มาตรฐาน กำหนดให้ตัวแปรตามคือ “2”

- สร้างสมการจำแนกกลุ่ม แสดงความสัมพันธ์ของตัวแปรตามกับพารามิเตอร์ในรูปแบบเชิงเส้นดังสมการที่ 2.13

#### ข) ขั้นตอนการทดสอบ (classifying phase)

ขั้นตอนนี้เป็นการทดสอบความแม่นยำของสมการการพยากรณ์จำแนกกลุ่ม เมื่อได้สมการ (discriminant function) จากขั้นตอนที่แล้ว ทำการทดสอบโดยการแทนค่าพารามิเตอร์ของกลุ่มข้อมูลชุดสำหรับการทดสอบลงในสมการจำแนกกลุ่มทั้ง 2 สมการ ผลลัพธ์ที่ได้จากสมการถูกนำมาใช้ในการจัดกลุ่ม จะได้ว่าถ้าผลลัพธ์ของสมการที่ 1 มีค่ามากกว่าผลลัพธ์ของสมการที่ 2 จะจัดให้ภาพตัวอย่างกึ่งนั้นอยู่ในกลุ่มที่ 1 กลุ่มที่มีการกำจัดเส้นดำไม่ผ่านเกณฑ์มาตรฐาน และถ้าผลลัพธ์ของสมการที่ 2 มีค่ามากกว่าผลลัพธ์ของสมการที่ 1 จะจัดให้ภาพตัวอย่างกึ่งนั้นอยู่ในกลุ่มที่ 2 กลุ่มที่มีการกำจัดเส้นดำผ่านเกณฑ์มาตรฐาน หรืออาจจะสรุปผลได้ว่าผลลัพธ์จากสมการใดมีค่ามากกว่ากึ่งในภาพตัวอย่างนั้นก็จัดอยู่ในกลุ่มนั้นด้วย

### 2) ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machines: SVM)

#### ก) ขั้นตอนการสอน (training phase)

ในงานวิจัยนี้เป็นการแบ่งกลุ่มข้อมูลแบบกลุ่มเดียว (one-class support vector classifier) โดยจะใช้กลุ่มข้อมูลชุดสำหรับการสอนในการสร้างโมเดล ดังนั้นจึงมีการวนกลุ่มชุดข้อมูลสำหรับการสอนดังอธิบายไว้ในขั้นตอนการเตรียมข้อมูล ซึ่งวิธี SVM มีขั้นตอนการสอนดังนี้

- นำค่าของกลุ่มมาตรฐานซึ่งอยู่ในรูปเวกเตอร์ทั้งสองกลุ่มมาพล็อตลงในพื้นที่คุณลักษณะ (feature space) จากนั้นทำการกำหนดสมการ เพื่อนำมาสร้างเป็นเส้นบน

ไฮเปอร์เพลนซึ่งแบ่งกลุ่มข้อมูลที่มีลักษณะเชิงเส้นสองกลุ่มออกจากกัน โดยกลุ่มข้อมูลทั้งสองฝั่งแทนด้วยค่า  $y$  มีสองค่าคือ  $y$  เท่ากับ 1 และ -1 เมื่ออินพุตข้อมูลเพื่อคำนวณหาค่า  $y$  ด้วยสมการที่ 2.17 ถ้าค่าดังกล่าวมีค่ามากกว่า 0 จะกำหนดค่า  $y$  ให้เท่ากับ 1 ดังสมการที่ 2.18 ซึ่งจะถูกจัดอยู่ในกลุ่มที่ 1 และถ้าค่าดังกล่าวมีค่าน้อยกว่า 0 จะกำหนดค่า  $y$  ให้เท่ากับ -1 ดังสมการที่ 2.19 ซึ่งจะถูกจัดให้อยู่ในกลุ่มที่ 2

- คำนวณเส้นแบ่งจากสมการที่ 2.21 จากนั้นนำค่าที่ได้ไปเขียนบนเส้นตรงตามแนวแกนโดยระยะทาง (maximum margin:  $d$ ) จากเส้นขอบไปยังไฮเปอร์เพลน

- เพิ่มเส้นขอบให้กับเส้นแบ่งทั้งสองข้างโดยเลือกจุดที่อยู่ใกล้เส้นตรง หากเส้นขอบของเส้นแบ่งใด ๆ มีความกว้างมากที่สุด แสดงให้เห็นว่าข้อมูลสองชุดมีการแยกกันชัดเจนมากที่สุด ดังนั้นเส้นแบ่งที่มีเส้นขอบกว้างที่สุดจึงเป็นเส้นแบ่งที่ดีที่สุด และเรียกตำแหน่งการสัมผัสข้อมูลที่ใกล้ที่สุดจากการเพิ่มขอบนี้ว่า “เวกเตอร์สนับสนุน” (Support Vector)

- เมื่อข้อมูลที่ใช้สำหรับการรู้จำมีการวางตัวกันในลักษณะตำแหน่งที่ไม่สามารถแบ่งกลุ่มได้โดยเส้นตรงตามวิธีการของซัพพอร์ตเวกเตอร์แมชชีนที่ไม่เป็นเส้นตรง จึงไม่สามารถนำมาใช้กับข้อมูลที่ไม่เป็นเชิงเส้นในลักษณะเช่นนี้ได้ จึงต้องใช้เคอร์เนลฟังก์ชัน (Kernel Function) มาช่วยโดยการเรียงตัวของข้อมูลใหม่ที่เรียกว่า “พื้นที่มิติสูง” (Higher Dimensional Space) โดยงานวิจัยได้ใช้เคอร์เนลฟังก์ชันเชิงเส้น และโพลิโนเมียล

โดยสร้างโมเดลการจำแนกกลุ่มระดับคุณภาพตามเกณฑ์มาตรฐานด้วยโปรแกรม WEKA เวอร์ชัน 3.8.1 ซึ่งมี 2 กลุ่ม คือ กลุ่มที่ 1 กลุ่มผ่านเกณฑ์มาตรฐาน และกลุ่มที่ 2 กลุ่มไม่ผ่านเกณฑ์มาตรฐาน โดยวิธีการนี้จะแสดงผลลัพธ์การวิเคราะห์ข้อมูลภาพเพื่อการจำแนกกลุ่มด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนด้วย ดังรูปที่ 3.17(ก) และ 3.17(ข) โดยส่วน Current Relation จะเป็นการสรุปรายละเอียดของข้อมูลภาพซึ่งจะแสดงรายละเอียดดังนี้

(1) Relation: ชื่อของของชุดข้อมูลที่โหลดเข้าไปใช้

(2) Instances: จำนวนแถวของข้อมูล

(3) Attributes: จำนวนแอตทริบิวต์

ส่วน Selected Attribute เป็นส่วนที่แสดงรายละเอียดของแต่ละแอตทริบิวต์ ซึ่งจะแสดงรายละเอียดดังนี้

(1) Name: ชื่อของแต่ละแอตทริบิวต์

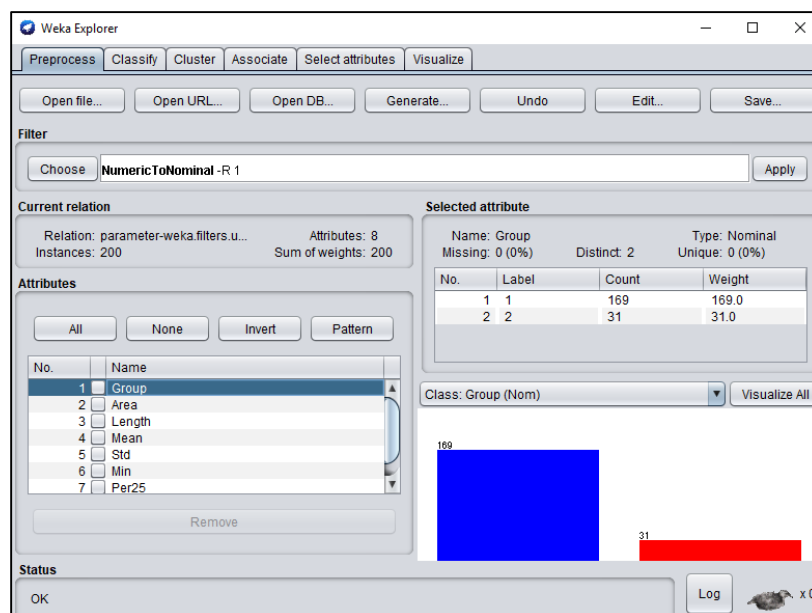
(2) Type: ลักษณะของข้อมูลในแอตทริบิวต์

(3) Missing: จำนวนของข้อมูลที่ขาดหายไป และเปอร์เซ็นต์ของข้อมูลที่หายไป

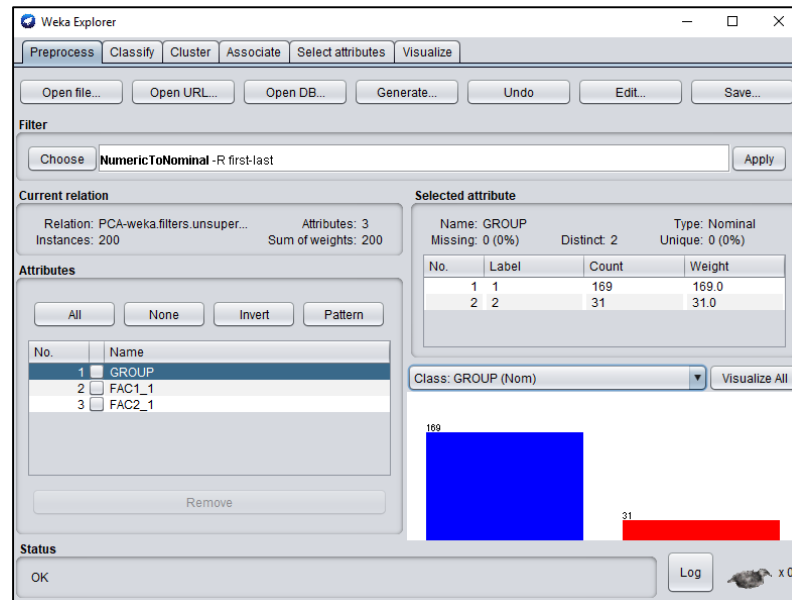
(4) Distinct: จำนวนของค่าของข้อมูลที่แตกต่างกันในแอตทริบิวต์

(5) Unique: จำนวนค่าของข้อมูลที่มีค่าไม่ซ้ำกันในแอตทริบิวต์

และอีกส่วนจะเป็นส่วนของตารางซึ่งจะแสดงรายละเอียดทางสถิติของข้อมูลในแอตทริบิวต์นั้น ประกอบด้วยค่าน้อยที่สุด (Minimum), ค่ามากที่สุด (Maximum), ค่าเฉลี่ย (Mean) และค่าเบี่ยงเบนมาตรฐาน (StdDev) ดังในรูปที่ 3.18(ก) และ 3.18(ข) แต่ข้อมูลส่วนของกลุ่มของคุณภาพที่ไม่ใช่ค่าตัวเลขก็จะแสดงเป็นความถี่ของข้อมูลแต่ละค่านั้นๆ ดังในรูปที่ 3.17(ก) และ 3.17(ข)



(ก)



(ข)

รูปที่ 3.17 ผลการวิเคราะห์เบื้องต้นด้วยโปรแกรม WEKA

(ก) กรณีที่ข้อมูลภาพไม่มีการลดขนาด (ข) กรณีที่ข้อมูลภาพมีการลดขนาด

Statistic		Value
Minimum		0
Maximum		6.4
Mean		3.081
StdDev		1.631

(ก)

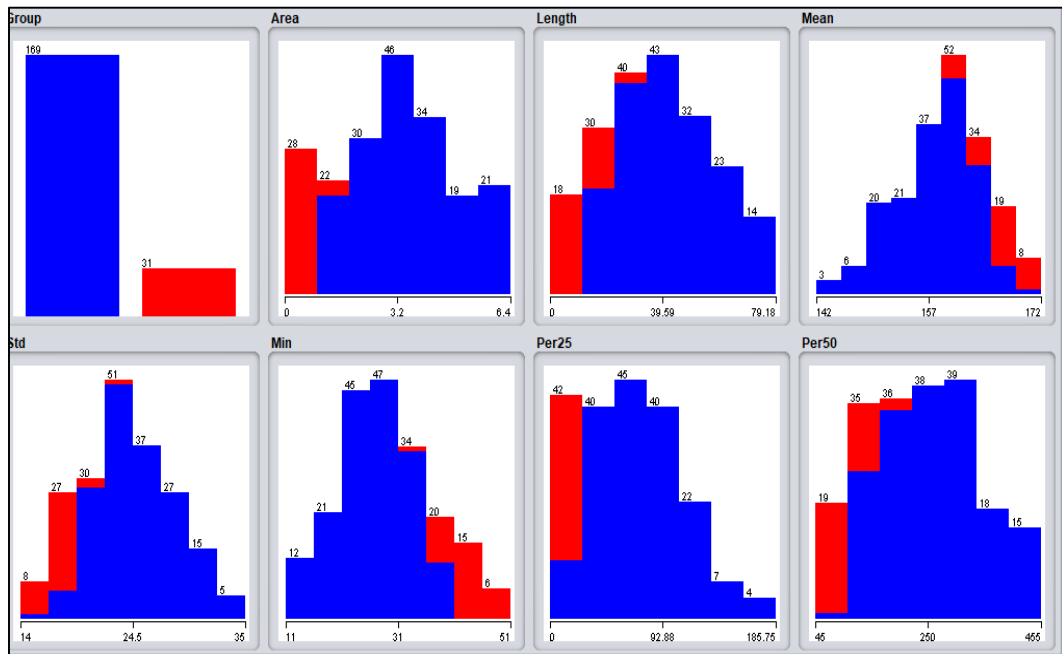
Statistic		Value
Minimum		-2.048
Maximum		3.702
Mean		0
StdDev		1

(ข)

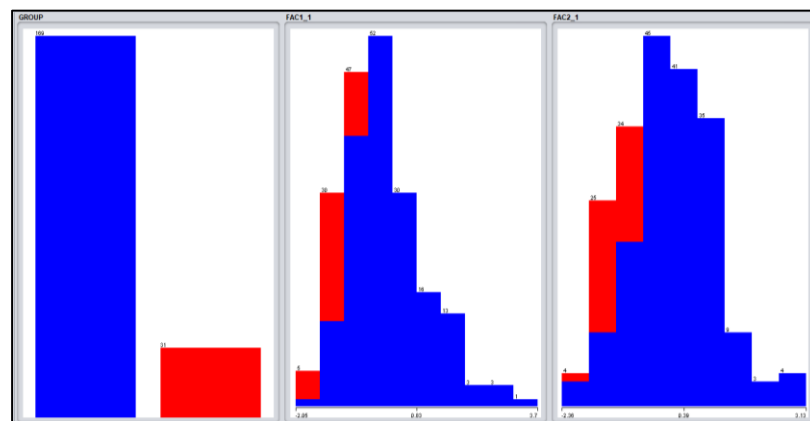
รูปที่ 3.18 รายละเอียดของแอดทริบิวต์ที่เป็นลักษณะของจำนวนตัวเลข

(ก) กรณีที่ข้อมูลภาพไม่มีการลดขนาด (ข) กรณีที่ข้อมูลภาพมีการลดขนาด

ส่วนสุดท้ายเป็นส่วนของการ visualization ซึ่งจะแสดงกราฟฮิสโทแกรม (histogram) ของแต่ละแอตทริบิวต์เมื่อเทียบกับคลาส (class) ของข้อมูล โดยที่สีฟ้า คือ คลาสของกลุ่มที่ 1 และ สีแดง คือ คลาสของกลุ่มที่ 2 ดังรูปที่ 3.21(ก) และ 3.21(ข)



(ก)



(ข)

รูปที่ 3.19 ฮิสโทแกรมแสดงความถี่ของทุกแอตทริบิวต์เทียบกับคลาส (class) ปลายทาง  
(ก) กรณีที่ข้อมูลภาพไม่มีการลดขนาด (ข) กรณีที่ข้อมูลภาพมีการลดขนาด



### ข) ขั้นตอนการทดสอบ (classifying phase)

หลังจากที่ผ่านกระบวนการสร้างเส้นแบ่งในขั้นตอนการสอนแล้ว ผลลัพธ์ที่ได้ออกมาคือโมเดลต้นแบบ เพื่อใช้สำหรับการแบ่งกลุ่มข้อมูลในขั้นตอนการทดสอบความแม่นยำของเส้นแบ่งหรือโมเดลต่อไป โดยจะใช้กลุ่มข้อมูลชุดสำหรับการทดสอบแบบวิธี Cross-validation Test ซึ่งจะมีการวนกลุ่มกันไปตามจำนวนกลุ่มข้อมูล ดังอธิบายไว้ในขั้นตอนการเตรียมข้อมูล

### 3.4 การประเมินผล (Evaluations)

ในงานวิจัยนี้มีการประเมินผลเพื่อวัดประสิทธิภาพและเปรียบเทียบผลลัพธ์การทำนายการจำแนกคุณภาพกุ้งจากปัจจัยคุณลักษณะซึ่งใช้โมเดลที่สร้างจากทั้ง 2 เทคนิค โดยใช้ค่าจากตารางการกระจายจำแนก (Classification distribution table) หรือเมทริกซ์ความสับสน (Confusion Matrix) ตามตารางที่ 3.2

ตารางที่ 3.4 Confusion matrix สำหรับการวัดประสิทธิภาพ

		Predicted Class	
		ไม่ผ่านมาตรฐาน	ผ่านมาตรฐาน
Actual Class	ไม่ผ่านมาตรฐาน	TN	FP
	ผ่านมาตรฐาน	FN	TP

โดยสามารถอธิบายค่าต่างๆใน Confusion Matrix ได้ดังนี้

True Positive (TP) หมายถึง กลุ่มเป้าหมายคือผ่านเกณฑ์ และตัวแบบทำนายว่าผ่านเกณฑ์

False Positive (FN) หมายถึง กลุ่มเป้าหมายคือผ่านเกณฑ์ แต่ตัวแบบทำนายว่าไม่ผ่านเกณฑ์

True Negative (TN) หมายถึง กลุ่มเป้าหมายคือไม่ผ่านเกณฑ์ และตัวแบบทำนายว่าไม่ผ่านเกณฑ์

False Positive (FP) หมายถึง กลุ่มเป้าหมายคือไม่ผ่านเกณฑ์ แต่ตัวแบบทำนายว่าผ่านเกณฑ์

นำมาพิจารณาจากค่าความถูกต้องต่างๆ เทคนิคการประเมินผลดังต่อไปนี้

1) ค่าความแม่นยำของโมเดล (Accuracy)

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}} \quad (3.9)$$

2) ค่าของโมเดลที่ทำนายได้ถูกต้องจากระดับคุณภาพ (Precision)

$$\text{Precision (Positive)} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (3.10)$$

$$\text{Precision (Negative)} = \frac{\text{FN}}{\text{TN} + \text{FN}} \quad (3.11)$$

3) ค่าของโมเดลที่ทำนายได้ (Recall)

$$\text{Recall (True)} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3.12)$$

$$\text{Recall (False)} = \frac{\text{FP}}{\text{TN} + \text{FP}} \quad (3.14)$$

4) ค่าเปรียบเทียบระหว่างค่า Precision และค่า Recall ของแต่ละกลุ่มเป้าหมาย

$$\text{F - Measure} = \frac{2 \times (\text{Recall} \times \text{Precision})}{\text{Recall} + \text{Precision}} \quad (3.15)$$

ในการพัฒนาเทคนิคการตรวจสอบข้อบกพร่องของกึ่งโดยประยุกต์ใช้เทคนิคประมวลผลภาพ มีขั้นตอนการดำเนินงานตามขั้นตอนต่างๆข้างต้น นอกจากนี้งานวิจัยนี้ยังทำการวัดประสิทธิภาพของการจำแนก เมื่อนำเทคนิคการตรวจสอบข้อบกพร่องที่พัฒนาขึ้นไปประยุกต์ใช้ พบว่าการตรวจสอบมีผลการดำเนินการวิจัยดังกล่าวสามารถอธิบายรายละเอียดไว้ในบทที่ 4

## บทที่ 4

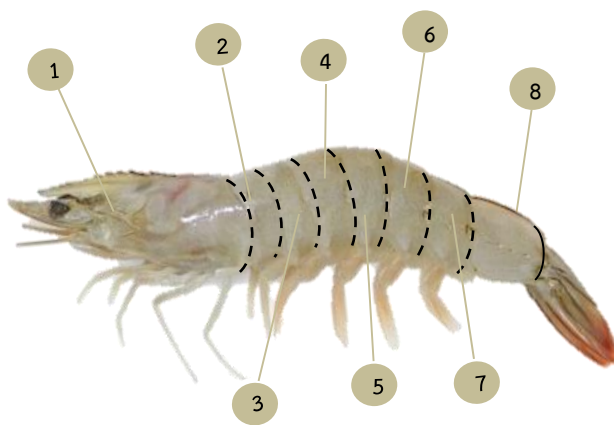
### ผลการดำเนินงาน

จากการศึกษาวิธีการสำหรับการพัฒนาเทคนิคการตรวจสอบเส้นด้ายซึ่งเป็นข้อบกพร่องของกึ่งแห้งแข็งในบทที่ 3 ได้ผลการวิเคราะห์ และผลการอภิปรายผลวิจัย ซึ่งเป็นผลจากการการวิเคราะห์เกณฑ์การจำแนก การเก็บข้อมูลภาพกึ่งเพื่อหาคุณลักษณะไปใช้เป็นพารามิเตอร์ในการจำแนกคุณภาพ การพัฒนาเทคนิคเพื่อจำแนกคุณภาพ และการประเมินประสิทธิภาพการจำแนก โดยผลการดำเนินการตามขั้นตอนวิธีการในการพัฒนาเทคนิคสามารถอธิบายรายละเอียดได้ดังนี้

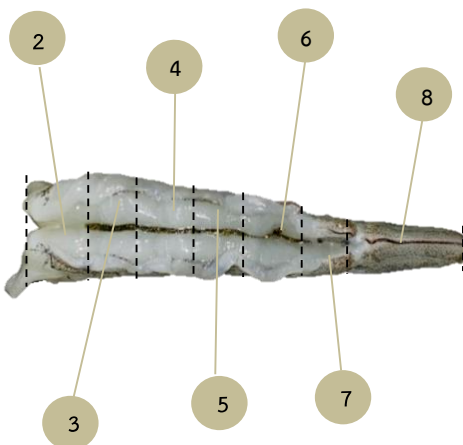
#### 4.1 ผลการวิเคราะห์เกณฑ์ในการจำแนก

กึ่งแวนดาไมซึ่งเป็นกึ่งที่ใช้ในงานวิจัย และเป็นสายพันธุ์กึ่งที่นิยมนำมาแปรรูปเป็นผลิตภัณฑ์แห้งแข็ง ประกอบด้วยส่วนหัว 1 ปล้อง ส่วนลำตัวมี 6 ปล้องหรือเรียกว่า 6 ข้อลำตัว และส่วนหาง 1 ปล้องดังรูปที่ 4.1(ก) จากเก็บข้อมูลวัดความยาวเฉลี่ยส่วนต่างๆของกึ่งที่ใช้ทำการวิจัยซึ่งเป็นกึ่งขนาด 50 ตัวต่อกิโลกรัมเทียบกับขนาดของปล้องพบว่า จากส่วนกริหัวถึงปลายกริหาง (ความยาวทั้งตัวกึ่ง) มีความยาวเฉลี่ย 12.2 เซนติเมตร จากโคนหัวถึงปลายกริหาง (ตั้งปล้องที่ 2 ถึงปลายหาง) มีความยาวเฉลี่ย 8.7 เซนติเมตร และเฉพาะแพนหางมีความยาวเฉลี่ย 1.4 เซนติเมตร และส่วนลำตัว (ตั้งแต่ต้นปล้องที่ 2 ถึงปลายปล้องที่ 7 มีความยาวเฉลี่ย 7.3 เซนติเมตร เมื่อผ่าบริเวณหลังกึ่งพบว่าเส้นด้ายซึ่งเป็นอวัยวะส่วนหนึ่งของระบบทางเดินอาหารมีความยาวตั้งแต่ส่วนต้นของปล้องที่ 2 ถึงส่วนกลางของปล้องที่ 8 แต่จากเกณฑ์มาตรฐานกำหนดว่าการกำจัดเส้นด้ายบริเวณหลังกึ่งจะไม่รวมเส้นด้ายที่อยู่ในบริเวณปล้องที่ 8 หรือปล้องที่ไม่มีการผ่า ดังนั้นความยาวของเส้นด้ายที่ตรวจสอบจะนับตั้งแต่บริเวณต้นปล้องที่ 2 จนถึงปลายปล้องที่ 7 ซึ่งวัดความยาวเฉลี่ยได้ 6.1 เซนติเมตร คิดเป็น 50% ของปล้องทั้งหมด หรือ 83.56% ของส่วนลำตัว จึงคาดคะเนความยาวของเส้นด้ายในแต่ละปล้องของลำตัว (ปล้องที่ 2 ถึง 6) ว่ามีปล้องละประมาณ 1 เซนติเมตร ดังนั้นจากเกณฑ์มาตรฐานที่กำหนดความยาวเส้นด้ายที่เป็นข้อบกพร่องต้องมีเส้นด้ายได้ไม่เกิน 1 ปล้องจึงประมาณว่ามีเส้นด้ายได้ไม่เกิน 1 เซนติเมตร แต่เนื่องด้วยกึ่งเป็นวัตถุดิบที่มีความไม่สม่ำเสมอของขนาดและสัดส่วนของความยาว จากข้อมูลข้างต้นจึงกำหนดให้ความยาวของเส้นด้ายที่ผ่านเกณฑ์มาตรฐานต้องมีค่าต่ำกว่า 1 เซนติเมตร เพื่อเป็นการลดข้อผิดพลาดจากการตรวจสอบและสร้างเกณฑ์ให้กับโปรแกรมสำหรับการจำแนกข้อบกพร่อง

จากมาตรฐานการจำแนกระดับกุ้งสดและแช่แข็งประเทศสหรัฐอเมริกา (United States Standards for Grades of Fresh and Frozen Shrimp) ใช้การคำนวณสัดส่วนเปอร์เซ็นต์ของน้ำหนักกุ้งที่ไม่ผ่านเกณฑ์มาตรฐานกับน้ำหนักของกุ้งทั้งหมดที่ได้จากการสุ่มตัวอย่าง พิจารณาเพื่อแบ่งเกรดของผลิตภัณฑ์กุ้งที่สามารถยอมรับได้ประกอบด้วย 2 เกรด คือ เกรดเอ และเกรดบี ซึ่งกำหนดว่าผลิตภัณฑ์กุ้งเกรดเอ (U.S. Grade A) มีจุดบกพร่องได้สูงสุด 15 จุด และผลิตภัณฑ์กุ้งเกรดบี (U.S. Grade B) มีจุดบกพร่องสูงสุด 30 จุด โดยจำนวนจุดบกพร่องสามารถเทียบกับเปอร์เซ็นต์น้ำหนักกุ้งที่ไม่ผ่านเกณฑ์มาตรฐานจากตารางข้อบกพร่องที่จะกำหนดเกณฑ์ตามขนาดกุ้ง ในวิจัยใช้ขนาดกุ้งที่เปอร์เซ็นต์ข้อบกพร่องต่ำที่สุด นั่นคือ กุ้งขนาดมากกว่า 40 ตัวต่อปอนด์เป็นเกณฑ์ซึ่งระบุว่าเปอร์เซ็นต์ของน้ำหนักกุ้งที่ไม่ผ่านเกณฑ์มาตรฐานที่ 1.00-3.00% คิดเป็นจุดบกพร่อง 3 จุด 3.01-5.00% คิดเป็นจุดบกพร่อง 8 จุด และ 5.01-10.00% คิดเป็นจุดบกพร่อง 16 จุด ดังนั้นเปอร์เซ็นต์ของน้ำหนักกุ้งที่ไม่ผ่านเกณฑ์มาตรฐานที่สามารถยอมรับได้ไม่เกิน 10.00%



(ก)



(ข)

รูปที่ 4.1 ลักษณะการแบ่งส่วนกุ้ง (ก) กุ้งขาว (ข) กุ้งผ่าหลัง

## 4.2 ผลการเก็บข้อมูลภาพกุ้ง

กุ้งได้ถูกถ่ายภาพภายในระบบดึงข้อมูลภาพที่ออกแบบไว้ดังแสดงในรูปที่ 3.5 สำหรับค่าเริ่มต้นของกล้องบันทึกภาพที่ใช้ในการถ่ายภาพเป็นดังตารางที่ 3.1 โดยเก็บตัวอย่างภาพกุ้งแต่ละระดับคุณภาพตามเกณฑ์มาตรฐานการนำเข้ากุ้งแช่แข็งของประเทศสหรัฐอเมริกา ซึ่งถูกแบ่งไว้ 2 ระดับ ประกอบด้วย ระดับผ่าน และระดับไม่ผ่านเกณฑ์มาตรฐาน การเก็บตัวอย่างภาพกุ้งในงานวิจัยนี้ตัวอย่างกุ้งจะต้องผ่านกระบวนการให้เป็นผลิตภัณฑ์ที่มีรูปแบบผ่าหลังกำจัดเส้นดำออก โดยเก็บภาพตัวอย่างกุ้งทั้งหมด 200 ตัวอย่าง ประกอบด้วยกุ้งที่ผ่านเกณฑ์มาตรฐานจำนวน 31 ตัวอย่างและกุ้งไม่ผ่านมาตรฐานจำนวน 169 ตัวอย่าง ซึ่งจำนวนตัวอย่างที่ใช้สำหรับการเก็บข้อมูลคิดจากจำนวนการสุ่มตัวแทนในสินค้าขนาดหนึ่งลีด สำหรับกุ้งที่มีขนาดตั้งแต่ 70 ตัวต่อปอนด์ขึ้นไปจะมีการสุ่มตัวแทนเพื่อตรวจสอบเกรด 2 ปอนด์ (0.91 กิโลกรัม) คิดเป็น 140 ตัวโดยประมาณ จึงเก็บภาพตัวอย่างกุ้งในจำนวนมากว่าเกณฑ์การสุ่มตัวแทน และมีเปอร์เซ็นต์ของน้ำหนักกุ้งที่ไม่ผ่านเกณฑ์มาตรฐานเกินจำนวนที่สามารถยอมรับได้

จากภาพกุ้งในแต่ละระดับมาตรฐาน พบว่าความยาวของเส้นดำบนข้อลำตัวของกุ้งส่งผลต่อการจำแนกกลุ่มมาตรฐาน โดยทั่วไปเส้นดำซึ่งเป็นอวัยวะส่วนลำไส้ของกุ้งมีสีค่อนข้างดำหรือมีบางส่วนที่อาจจะมีสีน้ำตาลหรือเทา ความแตกต่างของสีเส้นดำเกิดจากประเภทของอาหารที่กุ้งกิน จากการพิจารณาลักษณะของเส้นดำจะพบว่าความยาวของเส้นดำมีขนาดแตกต่างกันไปตามระดับมาตรฐาน กล่าวคือ กุ้งผ่านเกณฑ์มาตรฐานจะไม่ปรากฏเส้นดำซึ่งเป็นพื้นที่สีของสีดำ เทา หรือน้ำตาล หรือปรากฏเส้นดำประมาณไม่เกินหนึ่งข้อลำตัวซึ่งเป็นพื้นที่สีของสีดำ เทา หรือน้ำตาลบริเวณเนื้อกุ้งที่ถูกผ่าเล็กน้อย และกุ้งที่ไม่ผ่านเกณฑ์มาตรฐานจะปรากฏเส้นดำมากกว่าหนึ่งข้อลำตัวขึ้นไปซึ่งพื้นที่ของสีดำ เทา หรือน้ำตาลบริเวณเนื้อกุ้งที่ถูกผ่า ดังนั้นผู้วิจัยจึงใช้ลักษณะความยาว และความหนาพื้นที่สีในภาพระดับเทาที่ปรากฏบนตัวกุ้งที่มีการผ่าหลังเพื่อจำแนกระดับมาตรฐานของกุ้ง

## 4.3 ผลการพัฒนาเทคนิคการประมวลผลภาพสำหรับการตรวจสอบคุณภาพของกุ้ง

จากรายละเอียดในบทที่ 3 ซึ่งได้กล่าวถึงการพัฒนากระบวนการตรวจสอบคุณภาพด้วยการประยุกต์ใช้เทคนิคการประมวลผลภาพโดยใช้ซอฟต์แวร์ ImageJ เวอร์ชัน 1.51k โดยแบ่งกระบวนการออกเป็น 2 ส่วน ประกอบด้วย ส่วนแรกเป็นการประมวลผลภาพ เพื่อแยกคุณลักษณะที่

สำคัญ และส่วนที่สองเป็นกระบวนการจำแนกกลุ่มโดยใช้ข้อมูลคุณลักษณะที่ได้รับจากกระบวนการแรกซึ่งมีผลจากการศึกษาดังนี้

#### 4.3.1 การประมวลผลภาพ (image processing)

ขั้นตอนหลังจากการเก็บรวบรวมข้อมูลภาพถ่ายกึ่งภายในระบบดึงข้อมูลภาพ ประกอบด้วย การประมวลผลภาพเบื้องต้น การแยกบริเวณ และการหาคุณลักษณะ มีรายละเอียดดังต่อไปนี้

##### 4.3.1.1 ผลการประมวลผลภาพเบื้องต้น (pre-processing)

งานวิจัยนี้ใช้เทคนิคการแปลงภาพสีแบบ RGB เป็นภาพระดับเทาตามขั้นตอนในหัวข้อ 3.3.1.2 ภาพตัวอย่างกึ่งที่ได้จากการเก็บข้อมูลเป็นภาพที่อยู่ในระบบสีแบบ RGB ดังรูปที่ 4.2 (ก) โดยแต่ละพิกัดของภาพจะประกอบด้วยค่าของเซตความเข้มของค่าโทนแดง (R) ค่าโทนเขียว (G) และค่าโทนน้ำเงิน (B) จากนั้นแปลงให้เป็นภาพระดับเทาโดยใช้สมการที่ 3.1 ทำให้ภาพที่ได้สามารถนำมาประมวลผลได้ง่ายและรวดเร็วขึ้น นอกจากเพราะลดโทนสีที่ใช้ในจากประมวลผลแล้ว ดังรูปที่ 4.3(ข)



(ก)



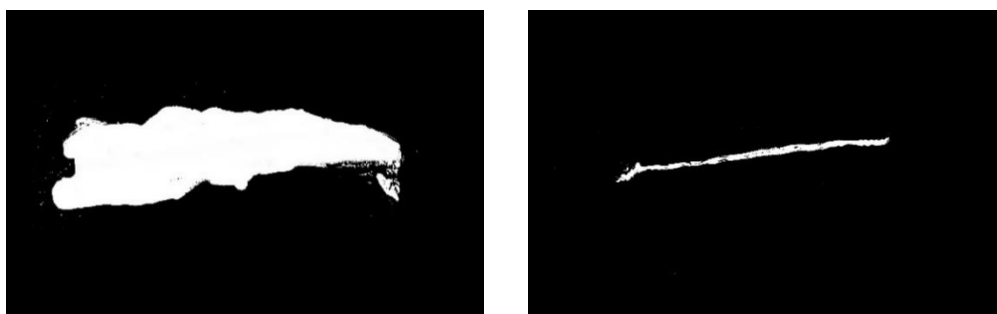
(ข)

รูปที่ 4.2 ภาพตัวอย่างกึ่ง (ก) ภาพระบบสี RGB (ข) ภาพระดับเทา

##### 4.3.1.2 ผลการแยกบริเวณ (image segmentation)

ขั้นตอนในการใช้ชุดคำสั่งเพื่อแยกบริเวณภาพที่สนใจออกจากส่วนอื่นในหัวข้อ 3.3.1.3 โดยใช้การปรับค่าขีดแบ่ง (thresholding value) ให้เหมาะสมตามฮิสโตแกรมของภาพ จากตัวอย่างของชุดคำสั่งในรูปที่ 3.8 นั่นคือปรับค่าขีดแบ่ง ส่วนตัวกึ่งให้ค่าอยู่ระหว่าง 83 ถึง

149 และส่วนเส้นดำปรับค่าขีดแบ่ง ให้อยู่ระหว่าง 83 ถึง 255 เมื่อใช้คำสั่งดังกล่าวในการแบ่งส่วนภาพด้วยการขีดแบ่ง จะได้รูปตัวอย่างกึ่งที่แยกบริเวณออกมาดังรูปที่ 4.3



(ก)

(ข)

รูปที่ 4.3 ภาพการแยกบริเวณด้วยเทคนิคเทรชโฮล (ก) ตัวกึ่ง (ข) เส้นดำกึ่ง

#### 4.3.1.3 ผลการสกัดคุณลักษณะ (feature extraction)

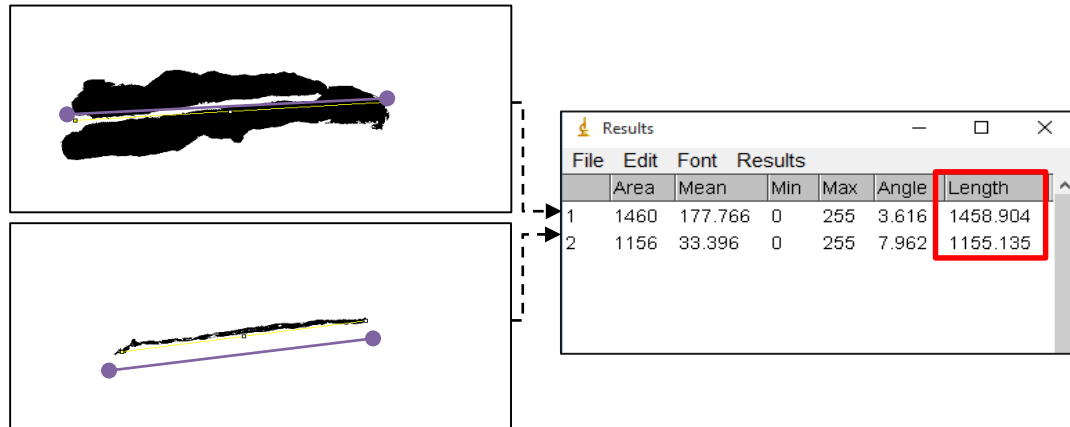
ผลของการสกัดคุณลักษณะที่คาดว่าจะมีส่วนต่อการจำแนกระดับคุณภาพของกึ่ง โดยรายละเอียดขั้นตอนได้แสดงในหัวข้อที่ 3.3.1.4 มีดังนี้

##### 1) คุณลักษณะทางรูปร่าง (shape features)

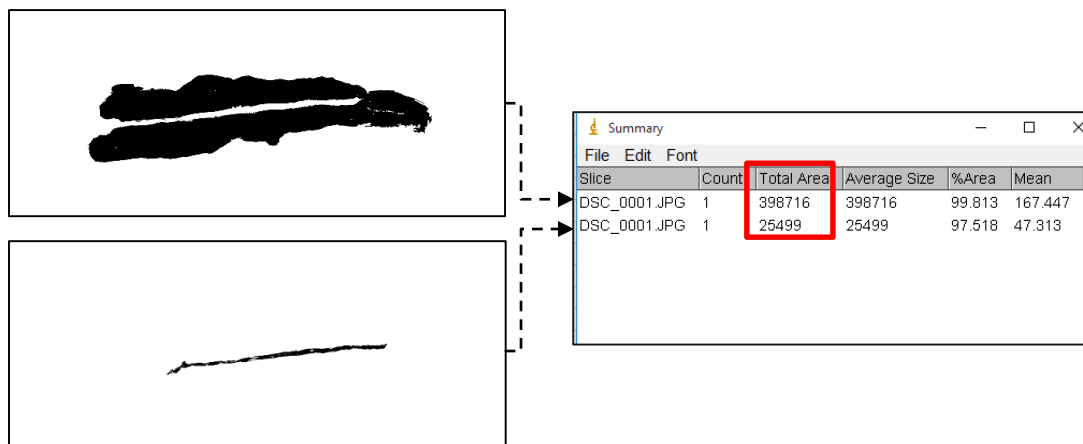
คุณลักษณะทางรูปร่างของเส้นดำที่สามารถดึงออกมาได้คือ ความยาวและพื้นที่ของเส้นดำ โดยภาพที่นำมาวัดความยาวหรือพื้นที่จะมีความหนาแน่นของจุดภาพหรือความละเอียดที่ไม่เท่ากัน โปรแกรมจะไม่สามารถระบุความยาวหรือพื้นที่ของภาพเป็นหน่วยตามระบบการวัดได้ แต่จะวัดขนาดของภาพนั้นในหน่วยพิกเซล (pixel) ซึ่งเป็นหน่วยพื้นฐานของภาพ โดยทั่วไปก่อนวัดความยาวหรือพื้นที่บนภาพด้วยชุดคำสั่งวัดในโปรแกรมจำเป็นต้องเทียบจุดภาพในหน่วยพิกเซล ให้เป็นหน่วยวัดตามระบบการวัดสากลก่อน จึงจะสามารถระบุขนาดที่แท้จริงของวัตถุบนภาพได้ แต่เนื่องจากในงานวิจัยนี้ใช้วิธีการเปรียบเทียบสัดส่วนระหว่างคุณลักษณะทางรูปร่างแต่ละคุณลักษณะของเส้นดำกับตัวกึ่งทั้งหมด เพื่อช่วยลดความคาดเคลื่อนจากขนาดของตัวกึ่งและเส้นดำที่ไม่สม่ำเสมอ จึงไม่จำเป็นต้องปรับเทียบเป็นหน่วยวัดตามระบบสากล ซึ่งทำให้ลดขั้นตอนและระยะเวลาในการวิเคราะห์ภาพถ่าย ซึ่งผลการแยกคุณลักษณะมีรายละเอียดดังนี้

ก) ความยาว (length) ผลการวัดจากเครื่องมือและชุดคำสั่งของโปรแกรม จะได้จำนวนจุดภาพซึ่งเป็นขนาดความยาวของเส้นดำที่ปรากฏ ( $L_v$ ) และขนาดความยาวของตัวกึ่งตั้งแต่ส่วนหัวถึงส่วนหาง ( $L_b$ ) ดังรูปที่ 4.4 จากนั้นนำไปคำนวณหาสัดส่วนด้วยสมการ 3.2

ข) พื้นที่ (area) ผลการวัดจากชุดคำสั่งของโปรแกรม จะได้ขนาดของพื้นที่จากการรวมแต่ละจุดภาพของบริเวณเส้นดำ ( $A_v$ ) และบริเวณตัวกึ่งทั้งหมด ( $A_b$ ) ดังรูปที่ 4.5 จากนั้นนำจำนวนของจุดภาพหรือพื้นที่บริเวณทั้งสองไปคำนวณหาสัดส่วนด้วยสมการ 3.3



รูปที่ 4.4 ผลการหาความยาวของเส้นดำและตัวกึ่ง



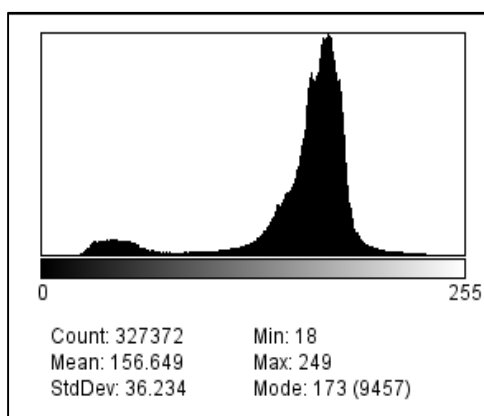
รูปที่ 4.5 ผลการหาพื้นที่ของเส้นดำและตัวกึ่ง

## 2) คุณลักษณะจากฮิสโตแกรม (histogram features)

ผลการดึงคุณลักษณะจากฮิสโตแกรมด้วยชุดคำสั่งของโปรแกรม พบว่าฮิสโตแกรมของภาพกึ่งที่มีลักษณะของเส้นดำอยู่ในกลุ่มผ่านมาตรฐานดังรูปที่ 3.12 มีความแตกต่างเมื่อเทียบกับฮิสโตแกรมของภาพกึ่งที่มีลักษณะของเส้นดำอยู่ในกลุ่มไม่ผ่านมาตรฐานดังรูปที่ 3.13 เช่นเดียวกับการทดลองเบื้องต้น (preliminary experiment) กับกลุ่มตัวอย่างทดสอบ ซึ่งคุณลักษณะของฮิสโตแกรมที่ดึงมาใช้ในการตรวจจับข้อบกพร่อง ได้แก่ ค่าเฉลี่ยของความสว่าง ค่าความสว่างที่ต่ำสุด ค่าเบี่ยงเบนมาตรฐาน และ จำนวนพิกเซลสะสม ณ เเปอร์เซ็นต์ไทล์ที่ 25, 50 และ 75 โดยค่า



คุณลักษณะของฮีสโตแกรมที่ได้จากการประมวลผลของโปรแกรมเป็นค่าของฮีสโตแกรมภาพ ซึ่งจะแสดงผลค่าต่างๆทางสถิติเป็นตำแหน่งของค่านั้น ยกตัวอย่างจากรูปที่ 4.6 จะได้ว่าค่า Mean หรือค่าเฉลี่ยของภาพเท่ากับ 156.649 ซึ่งเป็นค่าของตำแหน่งค่าเฉลี่ย เช่นเดียวกับค่า Min หรือค่าสว่างที่ต่ำที่สุด และค่าStdDev หรือค่าเบี่ยงเบนมาตรฐานเท่ากับ 18 และ 36.234 ซึ่งเป็นค่าของตำแหน่งความสว่างเช่นเดียวกัน



รูปที่ 4.6 ผลการประมวลผลฮีสโตแกรมภาพกึ่ง

ปัจจัยของคุณลักษณะซึ่งเป็นพารามิเตอร์ในการจำแนกจะมีจำนวนที่สกัดได้ไม่เท่ากันขึ้นอยู่กับลักษณะที่ส่งผลการตัดแยก ในการวิจัยจะประกอบด้วยคุณลักษณะหรือพารามิเตอร์ 8 ชนิดที่ได้จากขั้นตอนการสกัดคุณลักษณะนี้ โดยแต่ละพารามิเตอร์จะแสดงให้เห็นถึงความแตกต่างของกลุ่มคุณภาพทั้ง 2 กลุ่ม ซึ่งบางกรณีการใช้จำนวนพารามิเตอร์ที่มากจะทำให้โมเดลมีความแม่นยำมากขึ้น แต่จะทำให้ใช้เวลาในการสกัดคุณลักษณะและวิเคราะห์ผลมากขึ้นด้วย อย่างไรก็ตามการใช้จำนวนและลักษณะของพารามิเตอร์ที่ไม่เหมาะสมก็อาจจะทำให้ความแม่นยำลดลงหรือไม่เพิ่มขึ้น ทั้งนี้ขึ้นกับการพิจารณาและความเหมาะสมของการวิจัย

### 4.3.2 ผลการจำแนกระดับคุณภาพ (Quality Classification)

หลังจากได้ผลลัพธ์จากระบวนการประมวลผลภาพแล้ว นำข้อมูลที่ได้มาใช้เพื่อพัฒนาโมเดลสำหรับการจำแนกระดับคุณภาพ ซึ่งในกระบวนการนี้ประกอบด้วยขั้นตอน 2 ขั้นตอน นั่นคือ ขั้นตอนการเตรียมข้อมูล (preparing phase) และขั้นตอนการสร้างโมเดลเพื่อใช้ในการจำแนกระดับคุณภาพ (modeling phase) ซึ่งผลของกระบวนการนี้มีรายละเอียดแต่ละขั้นตอนดังนี้

#### 4.3.2.1 ผลการเตรียมข้อมูล (preparing phase)

ในส่วนนี้เป็นขั้นตอนการเตรียมข้อมูลให้เหมาะสมสำหรับการนำไปใช้กับโมเดลการจำแนก ซึ่งได้ผลจากการดำเนินการดังนี้

1) หลังจากประเมินคุณภาพของกึ่งในแต่ละภาพตัวอย่าง โดยระบุกลุ่มระดับคุณภาพตามเกณฑ์มาตรฐาน 2 กลุ่ม คือ กลุ่มผ่านเกณฑ์มาตรฐาน และกลุ่มไม่ผ่านเกณฑ์มาตรฐาน จากผลการจัดกลุ่ม สรุปได้ว่ามีภาพตัวอย่างกึ่งที่ผ่านเกณฑ์มาตรฐานการกำจัดเส้นดำจำนวน 31 ตัวอย่าง และภาพตัวอย่างกึ่งไม่ผ่านเกณฑ์มาตรฐานการกำจัดเส้นดำจำนวน 169 ตัวอย่าง

2) ผลจากการแบ่งกลุ่มข้อมูลสำหรับการสอนและทดสอบโมเดลจาก 3 วิธี มีรายละเอียดดังนี้

##### - วิธี self-consistency test

วิธีการแบ่งชุดข้อมูลนี้ได้นำมาใช้ในการจำแนกระดับคุณภาพด้วยวิธี discriminant analysis และซัพพอร์ตเวกเตอร์แมชชีน ชุดข้อมูลที่ใช้ในการสร้างโมเดลจำแนก และชุดข้อมูลที่ใช้ในการทดสอบประสิทธิภาพของโมเดลเป็นข้อมูลชุดเดียวกัน การแบ่งข้อมูลด้วยวิธีนี้จะให้ผลการวัดประสิทธิภาพที่มีค่าสูงมาก เนื่องจากเป็นข้อมูลชุดเดิมที่โมเดลได้รับการสอนมาแล้ว วิธีการนี้เหมาะสำหรับใช้ในการทดสอบประสิทธิภาพเพื่อดูแนวโน้มของโมเดลที่สร้างขึ้น ซึ่งถ้าความแม่นยำของโมเดลมีค่าน้อยแสดงว่าโมเดลไม่เหมาะสมกับข้อมูล จะไม่นำวิธีการจำแนกนั้นมาใช้ในการสร้างโมเดล

##### - วิธี split test

วิธีการแบ่งชุดข้อมูลนี้ได้นำมาใช้ในการจำแนกระดับคุณภาพวิธี discriminant analysis และซัพพอร์ตเวกเตอร์แมชชีนซึ่งแบ่งส่วนข้อมูลออกเป็น 70% ต่อ 30% โดยชุดข้อมูลส่วนที่หนึ่งจำนวน 140 ตัวอย่างใช้ในการสร้างโมเดลจำแนก และข้อมูลส่วนที่สองจำนวน 60 ตัวอย่างใช้ในการทดสอบประสิทธิภาพของโมเดล ซึ่งข้อดีของวิธีการแบ่งกลุ่มวิธีนี้คือใช้เวลาในการสร้างโมเดลน้อย และสะดวกต่อการแบ่งชุดข้อมูล แต่วิธีการนี้จะทำการสุ่มข้อมูลเพียงครั้งเดียว ซึ่งถ้าการสุ่มข้อมูลของชุดกลุ่มทดสอบมีลักษณะคล้ายหรือแตกต่างกับชุดกลุ่มข้อมูลที่ใช้สร้างโมเดลมาก อาจทำให้ผลการวัดประสิทธิภาพออกมาผิดพลาดได้

- วิธี leave-one-out cross-validation

รูปแบบการแบ่งชุดข้อมูลนี้ได้จะนำมาใช้ในการจำแนกระดับคุณภาพด้วยวิธี discriminant analysis และ วิธีซัพพอร์ตเวกเตอร์แมชชีน คือ ทำการแบ่งข้อมูลออกเป็น 2 ส่วน โดยที่ส่วนแรกมีจำนวน 199 ตัวอย่าง และส่วนที่สองมีจำนวน 1 ตัวอย่างจะใช้เป็นตัวอย่างทดสอบประสิทธิภาพของโมเดลทั้งหมด 200 ครั้ง ดังนี้

รอบที่ 1 ใช้ข้อมูลส่วนที่ 1 ทำการทดสอบ ใช้ส่วนที่เหลือสร้างโมเดล

รอบที่ 2 ใช้ข้อมูลส่วนที่ 2 ทำการทดสอบ ใช้ส่วนที่เหลือสร้างโมเดล

รอบที่ 3 ใช้ข้อมูลส่วนที่ 3 ทำการทดสอบ ใช้ส่วนที่เหลือสร้างโมเดล

⋮ ⋮

รอบที่ 199 ใช้ข้อมูลส่วนที่ 199 ทำการทดสอบ ใช้ส่วนที่เหลือสร้าง

โมเดล

รอบที่ 200 ใช้ข้อมูลส่วนที่ 200 ทำการทดสอบ ใช้ส่วนที่เหลือสร้าง

โมเดล

3) นำคุณลักษณะมาใช้เป็นพารามิเตอร์ของตัวจำแนก โดยนำคุณลักษณะทั้งหมดมาสังกัดให้เหลือเพียงชุดของคุณลักษณะที่สำคัญและน่าสนใจเท่านั้นด้วยการทดสอบทางสถิติ T-test จากกลุ่มตัวอย่างประชากร 2 กลุ่ม โดยนำค่าเฉลี่ยของแต่ละคุณลักษณะมาทำการทดสอบความแตกต่างของกลุ่มกึ่งผ่านและไม่ผ่านเกณฑ์มาตรฐานซึ่งก่อนนำมาทดสอบด้วย T-test ต้องทำการทดสอบเงื่อนไขความเท่ากันของความแปรปรวนด้วย F-test ที่ระดับความเชื่อมั่น 95% ( $\alpha = 0.05$ ) โดยการทดสอบด้วยโปรแกรม SPSS เวอร์ชัน 22 ได้แสดงตารางผลลัพธ์การวิเคราะห์ในภาคผนวก ข ซึ่งผลลัพธ์ที่ได้สามารถแปลผลได้ว่าคุณลักษณะของพื้นที่ (Area) ความยาว (Length) ค่าเฉลี่ย (Mean) ส่วนเบี่ยงเบนมาตรฐาน (Std) ค่าต่ำสุด (Min) เปอร์เซ็นไทล์ที่ 25 และ 50 (Per 25 และ 50) ค่าเฉลี่ยมีความแตกต่างอย่างมีนัยสำคัญ และสำหรับคุณลักษณะของ เปอร์เซ็นไทล์ที่ 75 (Per 75) ค่าเฉลี่ยของคุณลักษณะไม่มีความแตกต่างอย่างมีนัยสำคัญ ดังนั้นจึงสรุปได้ว่าคุณลักษณะที่จะนำมาใช้เป็นพารามิเตอร์ในการจำแนกกลุ่มมี 7 คุณลักษณะ คือ พื้นที่ (Area) ความยาว (Length) ค่าเฉลี่ย (Mean) ส่วนเบี่ยงเบนมาตรฐาน (Std) ค่าต่ำสุด (Min) เปอร์เซ็นไทล์ที่ 25 และ 50 (Per 25 และ 50)

4) เมื่อนำคุณลักษณะทั้ง 7 ชนิด มาลดขนาดข้อมูล (data reduction) ด้วยเทคนิค Principal Component Analysis (PCA) โดยใช้ข้อมูลคุณลักษณะจากตัวอย่างภาพกึ่ง

จำนวน 200 ภาพ จะได้ผลลัพธ์การวิเคราะห์ด้วยโปรแกรม SPSS เวอร์ชัน 22 ซึ่งได้แสดงผลลัพธ์การวิเคราะห์ในภาคผนวก ข สามารถสรุปผลได้ว่าการลดขนาดข้อมูลเทคนิคนี้เป็นวิธีการจัดกลุ่มปัจจัย (component) สามารถจัดกลุ่มปัจจัยได้มากที่สุดจำนวน 7 ปัจจัย ซึ่งเท่ากับจำนวนพารามิเตอร์ทั้งหมด โดยการจัดกลุ่มปัจจัยจะพิจารณาค่าไอเกนที่มากกว่า 0.7 สรุปผลการจัดกลุ่มพารามิเตอร์จาก 7 ปัจจัยสามารถจัดให้เหลือ 2 ปัจจัย สามารถอธิบายความแปรปรวนรวมสองปัจจัยเป็น 81.471 % เมื่อหมุนแกนปัจจัยโดยวิธี Varimax พบว่าค่าสัมประสิทธิ์สหสัมพันธ์สามารถจัดปัจจัยได้ ดังนี้

ปัจจัยที่ 1 ประกอบด้วยตัวแปร 5 พารามิเตอร์คือ ค่าเฉลี่ย (Mean), ส่วนเบี่ยงเบนมาตรฐาน (Std), พื้นที่ (Area), ค่าต่ำสุด (Min) และความยาว (Length)

ปัจจัยที่ 2 ประกอบด้วยตัวแปร 2 พารามิเตอร์คือ เปอร์เซนไทล์ที่ 25 (Per25) และเปอร์เซนไทล์ที่ 50 (Per50)

การลดขนาดข้อมูลให้เหลือ 2 ปัจจัย ถือว่าปัจจัยเป็นพารามิเตอร์ใหม่ ซึ่งค่าคะแนนปัจจัยหาได้จากสมการที่ 2.12 โดยจากตารางที่ ข-10 ในภาคผนวก ข สามารถคำนวณหาค่าคะแนนปัจจัยได้ดังนี้

ค่าคะแนนปัจจัยที่ 1 คือ

$$F_1 = -0.894Z_{11} + 0.844Z_{21} + 0.84Z_{31} - 0.79Z_{41} + 0.715Z_{51} + 0.26Z_{61} + 0.357Z_{71} \quad (4.1)$$

ค่าคะแนนปัจจัยที่ 2 คือ

$$F_2 = 0.398Z_{22} + 0.445Z_{32} - 0.429Z_{42} + 0.517Z_{52} + 0.881Z_{62} + 0.737Z_{72} \quad (4.2)$$

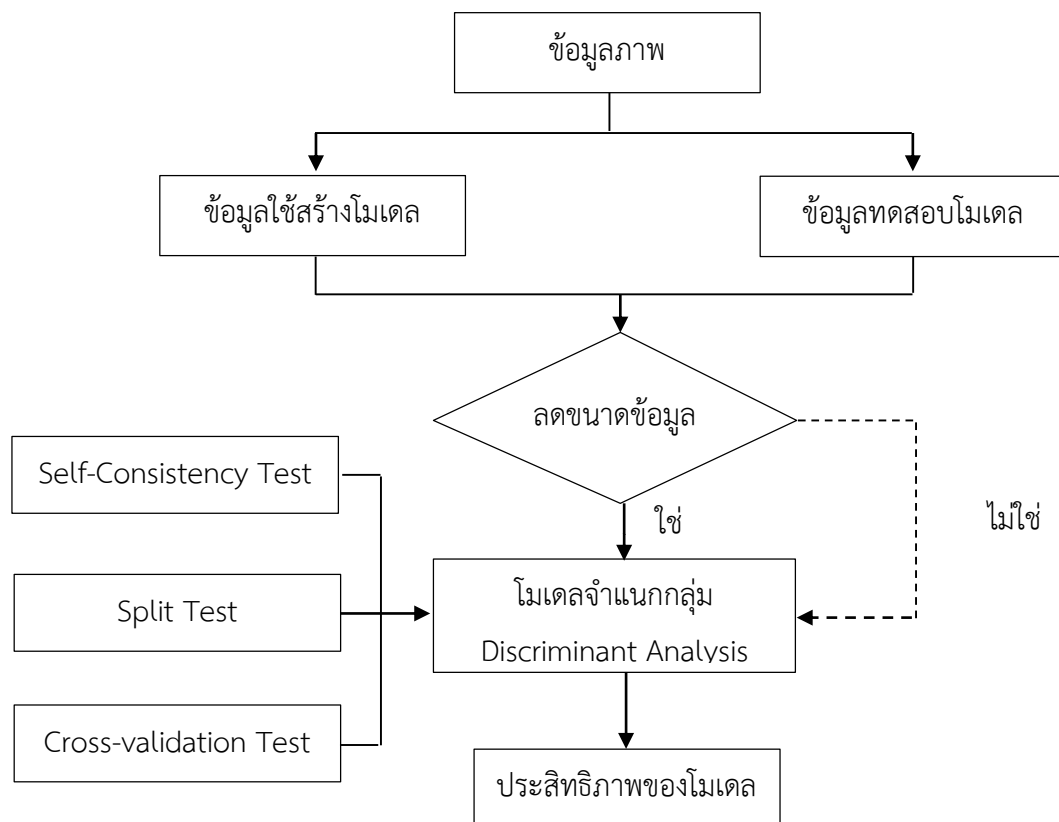
จากสมการการค่าคะแนนปัจจัยข้างต้น (สมการที่ 4.1 และ 4.2) โปรแกรม SPSS จะคำนวณค่าปัจจัยทั้ง 2 ปัจจัย ในตาราง ข-12 ภาคผนวก ข โดยใช้ชื่อคอลัมน์ว่า Factor1\_1 ซึ่งหมายถึง ปัจจัยที่ 1 ของการวิเคราะห์ครั้งที่ 1 และ Factor2\_1 หมายถึง ปัจจัยที่ 2 ของการวิเคราะห์ครั้งที่ 1

#### 4.3.2.2 ผลการสร้างโมเดล (modeling phase)

เทคนิคการลดขนาดข้อมูลถูกนำมาใช้ในการสร้างโมเดล ซึ่งโมเดลที่ได้จากขั้นตอนนี้เป็นตัวจำแนก (classifier) ในงานวิจัยนี้จะเปรียบเทียบผลความแม่นยำในการจำแนกระดับคุณภาพระหว่างการใช้เทคนิค 2 เทคนิค ประกอบด้วย วิธี discriminant analysis และวิธีซัพพอร์ตเวกเตอร์แมชชีน โดยมีรายละเอียดดังนี้

### 1) ผลการจำแนกกลุ่มด้วย Discriminant Analysis

วิธีนี้เป็นวิธีการทางสถิติวิธีหนึ่งที่ใช้เพื่อการจำแนกกลุ่ม ผู้วิจัยได้แบ่งข้อมูลออกเป็น 2 ชุด โดยชุดที่หนึ่งใช้สำหรับการสร้างโมเดล (training group) และชุดที่สองใช้สำหรับทดสอบประสิทธิภาพ (testing group) ขั้นตอนในกระบวนการจำแนกดังแผนผังรูปที่ 4.7 ข้อมูลภาพได้ถูกนำมาใช้วิเคราะห์ใน 2 กรณี คือ กรณีที่ข้อมูลภาพที่ไม่มีการลดขนาด และกรณีที่ข้อมูลภาพที่มีการลดขนาด หลังจากขั้นตอนการสร้างโมเดลเสร็จสิ้นแล้วโมเดลจะถูกนำมาทดสอบด้วยข้อมูลสำหรับทดสอบโมเดล ผลการวิเคราะห์ได้แสดงไว้ในภาคผนวก ค ซึ่งมีผลของแต่ละขั้นตอนการจำแนกกลุ่มดังนี้



รูปที่ 4.7 รูปแบบการวิเคราะห์การจำแนกกลุ่มด้วย discriminant analysis

#### ก) ขั้นตอนการสอน (training phase)

โดยขั้นตอนนี้ต้องตรวจสอบข้อสมมติฐานเบื้องต้นก่อนนำข้อมูลไปวิเคราะห์ประกอบด้วย 2 ประการ ดังนี้

- คุณลักษณะหรือพารามิเตอร์ทุกตัวต้องมีการแจกแจงแบบปกติตัวแปรหลายตัว พบว่าพารามิเตอร์ที่มีการแจกแจงแบบปกติตัวแปรหลายตัวของข้อมูลที่ไม่มีการลดขนาด คือพื้นที่ (Area) ความยาว (Length) ค่าเฉลี่ย (Mean) ส่วนเบี่ยงเบนมาตรฐาน (Std) และเปอร์เซ็นต์ไทล์ที่ 50 (Per 50) และข้อมูลที่มีการลดขนาด คือ สองพารามิเตอร์ที่ได้จากสมการค่าคะแนนปัจจัยจากสมการการลดข้อมูล

- เมทริกซ์ความแปรปรวนร่วมของพารามิเตอร์ในแต่ละกลุ่มเท่ากัน โดยใช้สถิติทดสอบ Box's M

สรุปได้ว่าข้อมูลเป็นไปตามข้อสมมติเบื้องต้นทั้ง 2 ข้อ ของการจำแนกประเภท จึงสามารถนำข้อมูลไปทำการวิเคราะห์จำแนกต่อไป

ผลการวิเคราะห์ค่าสัมประสิทธิ์ (ค่าถ่วงน้ำหนัก) และค่าคงที่ของสมการจำแนกกลุ่มโดยแยกเป็นกลุ่มตามวิธี Fisher's linear discrimination function จำนวนสมการจะมีเท่ากับจำนวนกลุ่มระดับคุณภาพตามเกณฑ์มาตรฐานในกรณีนี้มี 2 กลุ่ม คือ กลุ่มที่ 1 กลุ่มผ่านเกณฑ์มาตรฐาน และกลุ่มที่ 2 กลุ่มไม่ผ่านเกณฑ์มาตรฐาน ดังนั้นจึงมี 2 สมการ จากผลการวิเคราะห์จะได้สมการจำแนกของข้อมูลที่ไม่มีการลดขนาดและมีการลดขนาด ดังนี้

- สมการของข้อมูลที่ไม่มีการลดขนาด:

กลุ่มที่ 1 :

$$\hat{D}_1 = -1429.122 + 21.577\text{Area} - 0.426\text{Length} + 16.325\text{Mean} + 5.229\text{Std} + 0.318\text{Per50} \quad (4.3)$$

กลุ่มที่ 2 :

$$\hat{D}_2 = -1414.677 + 20.548\text{Area} - 0.477\text{Length} + 16.274\text{Mean} + 5.264\text{Std} + 0.300\text{Per50} \quad (4.4)$$

- สมการของข้อมูลที่มีการลดขนาด:

กลุ่มที่ 1 :

$$\hat{D}_1 = -0.798 + 0.311F_1 + 0.528F_2 \quad (4.5)$$

กลุ่มที่ 2 :

$$\hat{D}_2 = -3.515 - 2.562F_1 - 2.167F_2 \quad (4.6)$$

ผลการวิเคราะห์สามารถนำไปใช้พยากรณ์ข้อมูลใหม่ เพื่อให้ทราบว่าควรจัดให้อยู่กลุ่มใด โดยการแทนค่าพารามิเตอร์ลงในสมการทั้งสอง จากสมการของข้อมูลที่ไม่มีการลดขนาด พารามิเตอร์คือพื้นที่ (Area) ความยาว (Length) ค่าเฉลี่ย (Mean) ส่วนเบี่ยงเบนมาตรฐาน (Std) และเปอร์เซ็นต์ไทล์ที่ 50 (Per 50) ส่วนสมการของข้อมูลที่มีการลดขนาด พารามิเตอร์คือ ปัจจัยที่ 1 และปัจจัยที่ 2 ที่ได้จากสมการค่าคะแนนปัจจัยในขั้นตอนการลดขนาด

ข้อมูล ถ้าสมการของกลุ่มที่ 1 มีค่ามากกว่าสมการของกลุ่มที่ 2 จะจัดให้ภาพดังกล่าวอยู่ในกลุ่มที่ 1 แต่ถ้าสมการของกลุ่มที่ 2 มีค่ามากกว่าสมการของกลุ่มที่ 1 จะจัดให้ภาพดังกล่าวอยู่ในกลุ่มที่ 2 นั่นคือ ถ้าสมการใดมีค่ามากกว่าก็จัดอยู่ในกลุ่มนั้น

#### ข) ผลจากการทดสอบโมเดล (classifying phase)

ผลการทดสอบประสิทธิภาพความแม่นยำซึ่งแสดงในตารางที่ 4.1 และ 4.2 เป็นการแบ่งใช้ข้อมูลภาพ 2 กรณี คือ กรณีที่ข้อมูลภาพที่ไม่มีการลดขนาดจะใช้สมการจำแนกสมการที่ 4.3 และ 4.4 ส่วนกรณีที่ข้อมูลภาพที่มีการลดขนาดจะใช้สมการจำแนกสมการที่ 4.5 และ 4.6 ในแต่ละแบบมีลักษณะการแบ่งข้อมูลในการทดสอบประสิทธิภาพ 3 วิธี คือ วิธี self-consistency test, วิธี split test และวิธี leave-one-out cross-validation test โดยมีรายละเอียดดังต่อไปนี้

##### (1) กรณีที่ข้อมูลภาพที่ไม่มีการลดขนาด

- วิธี self-consistency test จะใช้ข้อมูลทั้งหมด 200 ตัวอย่าง ในการสร้างสมการและการพยากรณ์ พบว่า ในกลุ่มที่ 1 เดิมมีจำนวน 169 ตัวอย่าง จากการทำนายโดยใช้สมการจำแนกกลุ่มพบว่าทำนายได้ถูกต้องจำนวน 162 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 95.9 % ส่วนในกลุ่มที่ 2 เดิมมีจำนวน 31 ตัวอย่าง จากการทำนายโดยใช้สมการจำแนกกลุ่มพบว่าทำนายได้ถูกต้องจำนวน 31 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 100.0 % ดังนั้นสมการจำแนกกลุ่มสามารถจำแนกได้ถูกต้องจำนวน 193 ตัวอย่าง คิดเป็น 96.5 %

- วิธี split test จะใช้การทดสอบความแม่นยำของโมเดลด้วยสมการพยากรณ์จำแนกกลุ่มที่สร้างขึ้นจากชุดข้อมูลสำหรับสร้างโมเดลจำนวน 140 ตัวอย่าง และการทดสอบประสิทธิภาพความถูกต้องของสมการจำแนกกลุ่มด้วยการพยากรณ์ด้วยชุดข้อมูลสำหรับทดสอบโมเดลจำนวน 60 ตัวอย่าง พบว่าในกลุ่มที่ 1 เดิมมีจำนวน 54 ตัวอย่าง แต่จากการทำนายโดยใช้สมการจำแนกกลุ่มพบว่าทำนายได้ถูกต้องจำนวน 50 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 92.6 % ส่วนในกลุ่มที่ 2 เดิมมีจำนวน 6 ตัวอย่าง จากการทำนายโดยใช้สมการจำแนกกลุ่มพบว่าทำนายได้ถูกต้องจำนวน 6 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 100.0 % ดังนั้นสมการจำแนกกลุ่มสามารถจำแนกได้ถูกต้องจำนวน 56 ตัวอย่าง คิดเป็น 93.3%

-วิธี leave-one-out cross-validation test จะใช้ข้อมูลทั้งหมด 200 ตัวอย่าง ในการสร้างสมการและการพยากรณ์ พบว่า ในกลุ่มที่ 1 เดิมมีจำนวน 169 ตัวอย่าง แต่จากการทำนายโดยใช้สมการจำแนกกลุ่มพบว่าทำนายได้ถูกต้องจำนวน 160 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 94.7 % ส่วนในกลุ่มที่ 2 เดิมมีจำนวน 31 ตัวอย่าง จากการทำนายโดยใช้สมการ

จำแนกกลุ่มพบว่าทำนายได้ถูกต้องจำนวน 31 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 100.0 % ดังนั้นสมการจำแนกกลุ่มสามารถจำแนกได้ถูกต้องจำนวน 191 ตัวอย่าง คิดเป็น 95.5 %

(2) กรณีที่ข้อมูลภาพที่มีการลดขนาด

- วิธี self-consistency test จะใช้ข้อมูลทั้งหมด 200 ตัวอย่าง โดยเมื่อข้อมูลภาพที่มีการลดขนาดแล้วพารามิเตอร์ในการจำแนกจะเหลือ 2 พารามิเตอร์ ในการสร้างสมการและการพยากรณ์ พบว่า ในกลุ่มที่ 1 เดิมมีจำนวน 169 ตัวอย่าง แต่จากการทำนายโดยใช้สมการจำแนกกลุ่มพบว่าทำนายได้ถูกต้องจำนวน 155 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 91.7 % ส่วนในกลุ่มที่ 2 เดิมมีจำนวน 31 ตัวอย่าง จากการทำนายโดยใช้สมการจำแนกกลุ่มพบว่าทำนายได้ถูกต้องจำนวน 31 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 100.0 % ดังนั้นสมการจำแนกกลุ่มสามารถจำแนกได้ถูกต้องจำนวน 130 ตัวอย่าง คิดเป็น 93.0%

- วิธี split test จะใช้การทดสอบความแม่นยำของโมเดลด้วยสมการพยากรณ์จำแนกกลุ่มที่สร้างขึ้นจากชุดข้อมูลสำหรับสร้างโมเดลจำนวน 140 ตัวอย่าง และการทดสอบประสิทธิภาพความถูกต้องของสมการจำแนกกลุ่มด้วยการพยากรณ์ด้วยชุดข้อมูลสำหรับทดสอบโมเดลจำนวน 60 ตัวอย่าง พบว่าในกลุ่มที่ 1 เดิมมีจำนวน 54 ตัวอย่าง แต่จากการทำนายโดยใช้สมการจำแนกกลุ่มพบว่าทำนายได้ถูกต้องจำนวน 46 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 85.16 % ส่วนในกลุ่มที่ 2 เดิมมีจำนวน 6 ตัวอย่าง จากการทำนายโดยใช้สมการจำแนกกลุ่มพบว่าทำนายได้ถูกต้องจำนวน 6 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 100.0 % ดังนั้นสมการจำแนกกลุ่มสามารถจำแนกได้ถูกต้องจำนวน 52 ตัวอย่าง คิดเป็น 90.0 %

- วิธี leave-one-out cross-validation test จะใช้ข้อมูลทั้งหมด 200 ตัวอย่าง โดยเมื่อข้อมูลภาพที่มีการลดขนาดแล้วพารามิเตอร์ในการจำแนกจะเหลือ 2 พารามิเตอร์ ในการสร้างสมการและการพยากรณ์ พบว่า ในกลุ่มที่ 1 เดิมมีจำนวน 169 ตัวอย่าง แต่จากการทำนายโดยใช้สมการจำแนกกลุ่มพบว่าทำนายได้ถูกต้องจำนวน 155 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 91.7 % ส่วนในกลุ่มที่ 2 เดิมมีจำนวน 31 ตัวอย่าง จากการทำนายโดยใช้สมการจำแนกกลุ่มพบว่าทำนายได้ถูกต้องจำนวน 31 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 100.0 % ดังนั้นสมการจำแนกกลุ่มสามารถจำแนกได้ถูกต้องจำนวน 130 ตัวอย่าง คิดเป็น 93.0%



ตารางที่ 4.1 ผลลัพธ์การจำแนกกลุ่มด้วยข้อมูลภาพไม่มีการลดขนาด

	Group	Predicted Group Membership		Total	Accuracy	
		1	2			
Self-consistency	Actual Group Membership	1	162 (95.9%)	7 (4.1%)	169	96.5%
		2	0 (0%)	31 (100.0%)	31	
Split Test		1	50 (92.6%)	4 (7.4%)	54	93.3%
		2	0 (0%)	6 (100.0%)	6	
Leave-one-out		1	160 (94.7%)	9 (5.3%)	169	95.5%
		2	0 (0%)	31 (100.0%)	31	

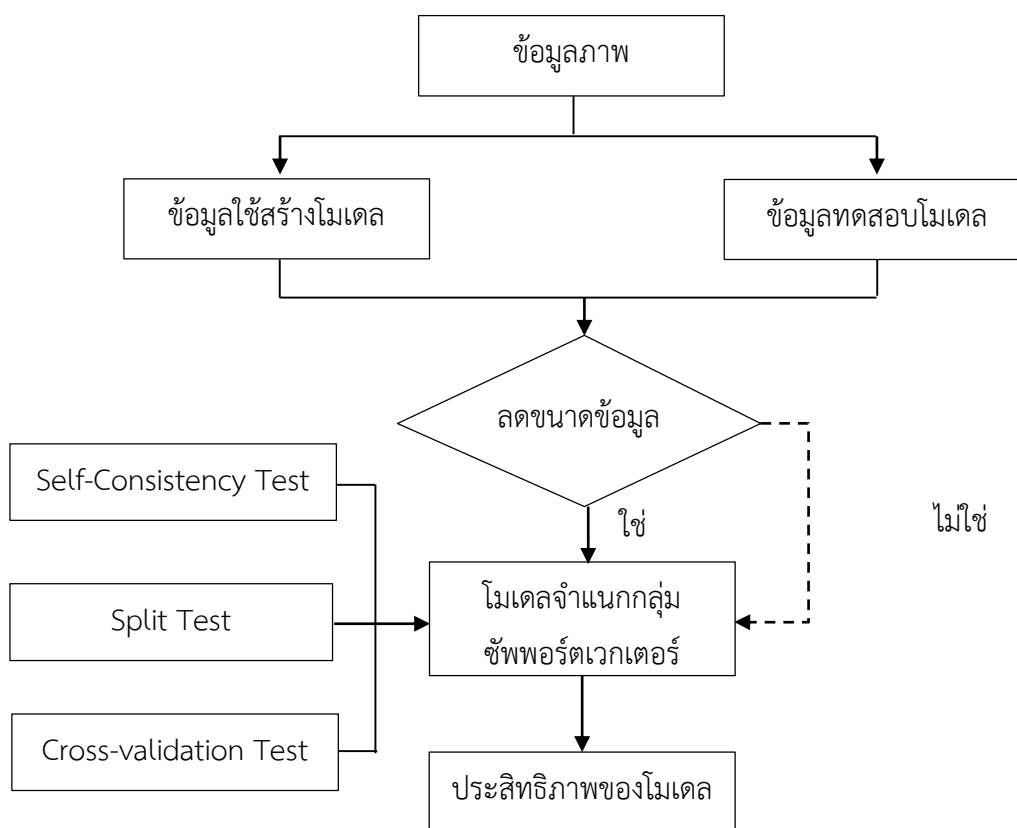
ตารางที่ 4.2 ผลลัพธ์การจำแนกกลุ่มด้วยข้อมูลภาพที่มีการลดขนาด

	Group	Predicted Group Membership		Total	Accuracy	
		1	2			
Self-consistency	Actual Group Membership	1	155 (91.7%)	14 (8.3%)	169	93.0%
		2	0 (0%)	31 (100.0%)	31	
Split Test		1	46 (85.16%)	8 (14.14%)	54	90.0%
		2	0 (0%)	6 (100.0%)	6	
Leave-one-out		1	155 (91.7%)	14 (8.3%)	169	93.0%
		2	0 (0%)	31 (100.0%)	31	

## 2) ผลการจำแนกด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีน

วิธีนี้เป็นการนำอัลกอริทึมมาใช้ในการวิเคราะห์ข้อมูลและจำแนกข้อมูล โดยการสร้างเส้นแบ่งแยกกลุ่มข้อมูลด้วยการหาสัมประสิทธิ์ของสมการที่ถูกป้อนเข้าสู่ขั้นตอนการสอนในการสร้างโมเดล โดยรูปแบบการวิเคราะห์จะแยกตามขั้นตอนในกระบวนการจำแนกดังแผนผังรูปที่ 4.9 การจำแนกระดับคุณภาพด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนจะใช้ข้อมูลภาพตัวอย่างทั้งหมด 200 ภาพ การทดสอบประสิทธิภาพการจำแนกข้อมูลภาพได้ใช้ค่าความแม่นยำเป็นตัวชี้วัดหลัก โดยข้อมูลภาพได้ถูกนำมาใช้วิเคราะห์ใน 2 กรณี คือ กรณีที่ข้อมูลภาพไม่มีการลดขนาด ซึ่งมีพารามิเตอร์สำหรับสร้างโมเดล 7 พารามิเตอร์ คือ พื้นที่ (Area) ความยาว (Length) ค่าเฉลี่ย (Mean) ส่วนเบี่ยงเบนมาตรฐาน (Std) ค่าต่ำสุด (Min) เปอร์เซ็นไทล์ที่ 25 และ 50 (Per 25 และ 50) และกรณีที่ข้อมูลภาพที่มีการลดขนาดแล้ว ซึ่งจะมีพารามิเตอร์สำหรับสร้างโมเดล 2 พารามิเตอร์ที่ได้จาก

สมการค่าคะแนนปัจจัยดังสมการที่ 4.1 และ 4.2 โดยโปรแกรม SPSS จะคำนวณค่าปัจจัยทั้ง 2 ปัจจัย ดังคอลัมน์ Factor1\_1 และ Factor2\_1 ในภาคผนวก ง จากนั้นจะนำข้อมูลสำหรับการสอนทั้ง 2 กรณีดังกล่าวเข้าสู่อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีนแบบหนึ่งต่อหนึ่ง (One-against-One) ด้วย ฟังก์ชันเคอร์เนลแบบเชิงเส้น (Linear Kernel) และแบบโพลิโนเมียล (Polynomial Kernel) เพื่อสร้างโมเดลที่ใช้ในการจำแนกข้อมูลภาพ ขั้นตอนสุดท้ายนำมาทดสอบโมเดลด้วยข้อมูลสำหรับทดสอบโมเดล มีขั้นตอนการดังนี้



รูปที่ 4.8 รูปแบบการวิเคราะห์การจำแนกกลุ่มด้วยซัพพอร์ตเวกเตอร์แมชชีน

#### ก) ขั้นตอนการสอน (training phase)

ในการสร้างโมเดลการจำแนกกลุ่มระดับคุณภาพตามเกณฑ์มาตรฐานในกรณีนี้มี 2 กลุ่ม คือ กลุ่มที่ 1 กลุ่มผ่านเกณฑ์มาตรฐาน และกลุ่มที่ 2 กลุ่มไม่ผ่านเกณฑ์มาตรฐาน โดยวิธีการนี้จะแสดงผลการวิเคราะห์ข้อมูลภาพเพื่อการจำแนกกลุ่มด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนด้วย จากโปรแกรม WEKA เวอร์ชัน 3.8.1 ผลการเรียนรู้และสร้างโมเดลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนจะแสดงวิธีการแบ่งข้อมูลในการทดสอบประสิทธิภาพ ฟังก์ชันเคอร์เนลที่ใช้ในการสร้างโมเดล คลาสการจำแนกจำนวน 2 คลาส ได้แก่ กลุ่มที่ 1 และกลุ่มที่ 2 สมการพร้อมค่าถ่วง

น้ำหนักของแอดทริบิวต์และระยะเวลาที่ใช้ในการสร้างโมเดล จากรูปที่ 4.9 และ 4.10 แสดงข้อมูล ทั้ง 2 กรณี คือ กรณีที่ข้อมูลภาพขนาดไม่มีการลดขนาด และกรณีที่ข้อมูลภาพที่มีการลดขนาดแล้วใช้วิธีการแบ่งข้อมูลในการทดสอบประสิทธิภาพด้วยวิธี self-consistency test ใช้ฟังก์ชันเคอร์เนลแบบเชิงเส้น (Linear Kernel) ในการสร้างโมเดล โดยสมการของแอดทริบิวต์กรณีที่ข้อมูลภาพไม่มีการลดขนาดประกอบด้วย 7 แอดทริบิวต์ ได้แก่ ความยาว (Length) พื้นที่ (Area) ค่าเฉลี่ย (Mean) ค่าเบี่ยงเบนมาตรฐาน (Std) ค่าความสว่างต่ำสุด (Min) เปอร์เซ็นไทล์ที่ 25 และ 50 (Per 25 และ 50) ใช้เวลา 0.06 วินาทีในการสร้างโมเดล ส่วนสมการของแอดทริบิวต์กรณีที่ข้อมูลภาพที่มีการลดขนาดประกอบด้วย 2 แอดทริบิวต์ ได้แก่ ปัจจัยที่ 1 (FAC1\_1) และ ปัจจัยที่ 2 (FAC2\_1) ใช้เวลา 0.01 วินาทีในการสร้างโมเดล

ส่วนผลการเรียนรู้และสร้างโมเดลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีน จากโปรแกรม WEKA ด้วยลักษณะข้อมูล วิธีการแบ่งข้อมูลการทดสอบประสิทธิภาพ และฟังก์ชันเคอร์เนลอื่นๆได้แสดงไว้ในภาคผนวก ง

```

Test mode:  evaluate on training data
=== Classifier model (full training set) ===
Kernel used:  Linear Kernel: K(x,y) = <x,y>
Classifier for classes:  1, 2
Machine linear:  showing attribute weights, not support vectors.
    -1.7423 * (normalized) Area
+   -1.1549 * (normalized) Length
+    0.8161 * (normalized) Mean
+   -0.8634 * (normalized) Std
+    2.3102 * (normalized) Min
+   -2.3229 * (normalized) Per25
+   -1.5563 * (normalized) Per50
-    0.5219
Time taken to build model: 0.06 seconds

```

รูปที่ 4.9 ผลการสร้างโมเดลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนกรณีที่ข้อมูลภาพไม่มีการลดขนาด

```

Test mode:  evaluate on training data
=== Classifier model (full training set) ===
Kernel used: Linear Kernel: K(x,y) = <x,y>
Classifier for classes: 1, 2
Machine linear: showing attribute weights, not support vectors.
      -3.8359 * (normalized) FAC1_1
+     -4.1539 * (normalized) FAC2_1
+       1.7109
Time taken to build model: 0.01 seconds

```

รูปที่ 4.10 ผลการสร้างโมเดลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนกรณีที่มีข้อมูลภาพที่มีการลดขนาด

ข) ผลจากการทดสอบโมเดล (classifying phase)

ผลการวัดประสิทธิภาพของโมเดลในการจำแนกที่สร้างขึ้นมาพบว่ามีค่าที่สำคัญอยู่ 5 ค่า นั่นคือ ค่าเฉลี่ยความคลาดเคลื่อนสมบูรณ์ (Mean Absolute Error: MAE) ความคลาดเคลื่อนเฉลี่ยกำลังสอง (Root Mean Square Error: RMSE) ซึ่งทั้ง 2 กรณีมีค่า MAE และ RMSE น้อยทั้ง 2 ค่า นั่นแสดงว่าโมเดลทั้งสองสามารถประมาณค่าได้ใกล้เคียงกับค่าจริง ค่าความถูกต้องในการจำแนก (Correctly Classified Instances) กรณีแรกมีค่าเท่ากับ 99.5% แสดงว่าจากข้อมูลภาพจำนวน 200 ตัวอย่าง มีการทำนายข้อมูลถูกต้อง 199 ตัวอย่าง หรือคิดเป็น 99.5% ของข้อมูลภาพทั้งหมด กรณีที่สองมีค่าเท่ากับ 98.0% แสดงว่าจากข้อมูลภาพจำนวน 200 ตัวอย่าง มีการทำนายข้อมูลถูกต้อง 196 ตัวอย่าง หรือคิดเป็น 98.0% ของข้อมูลภาพทั้งหมด ค่าความผิดพลาดในการทำนาย (Incorrectly Classified Instances) แสดงว่าจากข้อมูลภาพจำนวน 200 ตัวอย่าง กรณีแรกมีการทำนายข้อมูลไม่ถูกต้อง 1 ตัวอย่าง หรือคิดเป็น 0.5% กรณีที่สองมีการทำนายข้อมูลไม่ถูกต้อง 4 ตัวอย่าง หรือคิดเป็น 2.0% และสุดท้ายเมทริกซ์ความสับสน (Confusion Matrix) เป็นตารางที่แนวคอลัมน์จะบ่งบอกถึงค่าที่ได้ทำนายออกมาส่วนแนวนอนจะเป็นค่าของข้อมูลจริง ดังแสดงในรูปที่ 4.11 และ 4.12

```

=== Evaluation on training set ===
Time taken to test model on training data: 0 seconds
Correctly Classified Instances      199      99.5  %
Incorrectly Classified Instances    1        0.5  %
Kappa statistic                    0.9812
Mean absolute error                0.005
Root mean squared error            0.0707
Relative absolute error            1.8917 %
Root relative squared error        19.5376 %
Total Number of Instances         200

=== Confusion Matrix ===
  a  b  <-- classified as
168  1 | a = 1
  0 31 | b = 2

```

รูปที่ 4.11 ผลการทดสอบโมเดลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนกรณีข้อมูลภาพไม่มีการลดขนาด

```

=== Evaluation on training set ===
Time taken to test model on training data: 0.01 seconds
Correctly Classified Instances      196      98  %
Incorrectly Classified Instances    4        2  %
Kappa statistic                    0.9194
Mean absolute error                0.02
Root mean squared error            0.1414
Relative absolute error            7.567 %
Root relative squared error        39.0752 %
Total Number of Instances         200

=== Confusion Matrix ===
  a  b  <-- classified as
169  0 | a = 1
  4 27 | b = 2

```

รูปที่ 4.12 ผลการทดสอบโมเดลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนกรณีข้อมูลภาพมีการลดขนาด

ผลการทดสอบประสิทธิภาพของโมเดลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนมีการแบ่งทดสอบเป็นหลายกรณีดังกล่าวไว้ข้างต้น ผลการทดสอบโมเดลจากโปรแกรม WEKA กรณีอื่นจึงแสดงไว้ในภาคผนวก ง ดังนั้นสามารถสรุปผลการทดสอบประสิทธิภาพความแม่นยำของวิธีซัพพอร์ตเวกเตอร์แมชชีนในแต่ละกรณีดังตารางที่ 4.3 และ 4.4 โดยมีรายละเอียดดังนี้

(1) กรณีที่ข้อมูลภาพที่ไม่มีการลดขนาด

(1.1) ฟังก์ชันเคอร์เนลแบบเชิงเส้น (Linear Kernel)

- วิธี self-consistency test ได้ใช้ข้อมูลทั้งหมด 200 ตัวอย่าง ในการสร้างโมเดลและการพยากรณ์ พบว่า ในกลุ่มที่ 1 เดิมมีจำนวน 169 ตัวอย่าง แต่จากการทำนายพบว่าทำนายได้ถูกต้องจำนวน 169 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 100.0 % ส่วนในกลุ่มที่ 2 เดิมมีจำนวน 31 ตัวอย่าง จากการทำนายพบว่าทำนายได้ถูกต้องจำนวน 31 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 100.0 % ดังนั้นโมเดลสามารถจำแนกได้ถูกต้องจำนวน 200 ตัวอย่าง คิดเป็น 100.0 %

- วิธี split test ได้ใช้การทดสอบความแม่นยำของโมเดลที่สร้างขึ้นจากชุดข้อมูลสำหรับสอนจำนวน 140 ตัวอย่าง และการทดสอบประสิทธิภาพความถูกต้องด้วยการพยากรณ์ด้วยชุดข้อมูลสำหรับทดสอบโมเดลจำนวน 60 ตัวอย่าง พบว่าในกลุ่มที่ 1 เดิมมีจำนวน 48 ตัวอย่าง แต่จากการทำนายพบว่าทำนายได้ถูกต้องจำนวน 48 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 100.0 % ส่วนในกลุ่มที่ 2 เดิมมีจำนวน 12 ตัวอย่าง จากการทำนายพบว่าทำนายได้ถูกต้องจำนวน 8 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 66.7 % ดังนั้นโมเดลสามารถจำแนกได้ถูกต้องจำนวน 56 ตัวอย่าง คิดเป็น 93.3%

- วิธี leave-one-out cross-validation test ได้ใช้ข้อมูลในการพยากรณ์จำนวน 200 ตัวอย่าง พบว่า ในกลุ่มที่ 1 เดิมมีจำนวน 169 ตัวอย่าง แต่จากการทำนายพบว่าทำนายได้ถูกต้องจำนวน 169 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 100.0 % ส่วนในกลุ่มที่ 2 เดิมมีจำนวน 31 ตัวอย่าง จากการทำนายโดยใช้สมการจำแนกกลุ่มพบว่าทำนายได้ถูกต้องจำนวน 21 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 67.7% ดังนั้นสมการจำแนกกลุ่มสามารถจำแนกได้ถูกต้องจำนวน 190 ตัวอย่าง คิดเป็น 90.0 %

(1.2) ฟังก์ชันเคอร์เนลแบบโพลีโนเมียล (Polynomial Kernel)

- วิธี self-consistency test ได้ใช้ข้อมูลทั้งหมด 200 ตัวอย่าง ในการสร้างโมเดลและการพยากรณ์ พบว่า ในกลุ่มที่ 1 เดิมมีจำนวน 169 ตัวอย่าง แต่จากการทำนายพบว่าทำนายได้ถูกต้องจำนวน 169 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 100.0 % ส่วนในกลุ่มที่ 2 เดิมมีจำนวน 31 ตัวอย่าง จากการทำนายพบว่าทำนายได้ถูกต้องจำนวน 31 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 100.0 % ดังนั้นโมเดลสามารถจำแนกได้ถูกต้องจำนวน 200 ตัวอย่าง คิดเป็น 100.0 %

- วิธี split test ได้ใช้การทดสอบความแม่นยำของโมเดลที่สร้างขึ้นจากชุดข้อมูลสำหรับสอนจำนวน 140 ตัวอย่าง และการทดสอบประสิทธิภาพความถูกต้องด้วยการพยากรณ์ด้วยชุดข้อมูลสำหรับทดสอบโมเดลจำนวน 60 ตัวอย่าง พบว่าในกลุ่มที่ 1 เดิมมีจำนวน 48 ตัวอย่าง แต่จากการทำนายพบว่าทำนายได้ถูกต้องจำนวน 48 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 100.0 % ส่วนในกลุ่มที่ 2 เดิมมีจำนวน 12 ตัวอย่าง จากการทำนายพบว่าทำนายได้ถูกต้องจำนวน 0 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 0.0 % ดังนั้นโมเดลสามารถจำแนกได้ถูกต้องจำนวน 48 ตัวอย่าง คิดเป็น 80.0 %

- วิธี leave-one-out cross-validation test ใช้ข้อมูลในการพยากรณ์จำนวน 200 ตัวอย่าง พบว่า ในกลุ่มที่ 1 เดิมมีจำนวน 169 ตัวอย่าง แต่จากการทำนายพบว่าทำนายได้ถูกต้องจำนวน 169 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 100.0 % ส่วนในกลุ่มที่ 2 เดิมมีจำนวน 31 ตัวอย่าง จากการทำนายโดยใช้สมการจำแนกกลุ่มพบว่าทำนายได้ถูกต้องจำนวน 0 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 0.0 % ดังนั้นสมการจำแนกกลุ่มสามารถจำแนกได้ถูกต้องจำนวน 169 ตัวอย่าง คิดเป็น 84.5 %

## (2) กรณีที่ข้อมูลภาพที่มีการลดขนาด

### (2.1) ฟังก์ชันเคอร์เนลแบบเชิงเส้น (Linear Kernel)

- วิธี self-consistency test ได้ใช้ข้อมูลทั้งหมด 200 ตัวอย่าง ในการสร้างโมเดลและการพยากรณ์ พบว่า ในกลุ่มที่ 1 เดิมมีจำนวน 169 ตัวอย่าง แต่จากการทำนายพบว่าทำนายได้ถูกต้องจำนวน 169 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 100.0 % ส่วนในกลุ่มที่ 2 เดิมมีจำนวน 31 ตัวอย่าง จากการทำนายพบว่าทำนายได้ถูกต้องจำนวน 27 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 87.1 % ดังนั้นโมเดลสามารถจำแนกได้ถูกต้องจำนวน 196 ตัวอย่าง คิดเป็น 98.0 %

- วิธี split test ได้ใช้การทดสอบความแม่นยำของโมเดลที่สร้างขึ้นจากชุดข้อมูลสำหรับสอนจำนวน 140 ตัวอย่าง และการทดสอบประสิทธิภาพความถูกต้องด้วยการพยากรณ์ด้วยชุดข้อมูลสำหรับทดสอบโมเดลจำนวน 60 ตัวอย่าง พบว่าในกลุ่มที่ 1 เดิมมีจำนวน 48 ตัวอย่าง แต่จากการทำนายพบว่าทำนายได้ถูกต้องจำนวน 48 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 100.0 % ส่วนในกลุ่มที่ 2 เดิมมีจำนวน 12 ตัวอย่าง จากการทำนายพบว่าทำนายได้ถูกต้องจำนวน 0 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 0.0 % ดังนั้นโมเดลสามารถจำแนกได้ถูกต้องจำนวน 48 ตัวอย่าง คิดเป็น 80.0 %

- วิธี leave-one-out cross-validation test ใช้ข้อมูลในการพยากรณ์จำนวน 200 ตัวอย่าง พบว่า ในกลุ่มที่ 1 เดิมมีจำนวน 169 ตัวอย่าง แต่จากการทำนายพบว่าทำนายได้ถูกต้องจำนวน 169 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 100.0 % ส่วนในกลุ่มที่ 2 เดิม

มีจำนวน 31 ตัวอย่าง จากการทำนายโดยใช้สมการจำแนกกลุ่มพบว่าทำนายได้ถูกต้องจำนวน 19 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 54.8 % ดังนั้นสมการจำแนกกลุ่มสามารถจำแนกได้ถูกต้องจำนวน 188 ตัวอย่าง คิดเป็น 94.0 %

(2.2) ฟังก์ชันเคอร์เนลแบบโพลีโนเมียล (Polynomial Kernel)

- วิธี self-consistency test ได้ใช้ข้อมูลทั้งหมด 200 ตัวอย่าง ในการสร้างโมเดลและการพยากรณ์ พบว่า ในกลุ่มที่ 1 เดิมมีจำนวน 169 ตัวอย่าง แต่จากการทำนายพบว่าทำนายได้ถูกต้องจำนวน 169 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 100.0 % ส่วนในกลุ่มที่ 2 เดิมมีจำนวน 31 ตัวอย่าง จากการทำนายพบว่าทำนายได้ถูกต้องจำนวน 0 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 0.0 % ดังนั้นโมเดลสามารถจำแนกได้ถูกต้องจำนวน 169 ตัวอย่าง คิดเป็น 84.5 %

- วิธี split test ได้ใช้การทดสอบความแม่นยำของโมเดลที่สร้างขึ้นจากชุดข้อมูลสำหรับสอนจำนวน 140 ตัวอย่าง และการทดสอบประสิทธิภาพความถูกต้องด้วยการพยากรณ์ด้วยชุดข้อมูลสำหรับทดสอบโมเดลจำนวน 60 ตัวอย่าง พบว่าในกลุ่มที่ 1 เดิมมีจำนวน 48 ตัวอย่าง แต่จากการทำนายพบว่าทำนายได้ถูกต้องจำนวน 48 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 100.0 % ส่วนในกลุ่มที่ 2 เดิมมีจำนวน 12 ตัวอย่าง จากการทำนายพบว่าทำนายได้ถูกต้องจำนวน 0 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 0.0 % ดังนั้นโมเดลสามารถจำแนกได้ถูกต้องจำนวน 48 ตัวอย่าง คิดเป็น 80.0 %

- วิธี leave-one-out cross-validation test ใช้ข้อมูลในการพยากรณ์จำนวน 200 ตัวอย่าง พบว่า ในกลุ่มที่ 1 เดิมมีจำนวน 169 ตัวอย่าง แต่จากการทำนายพบว่าทำนายได้ถูกต้องจำนวน 169 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 100.0 % ส่วนในกลุ่มที่ 2 เดิมมีจำนวน 31 ตัวอย่าง จากการทำนายโดยใช้สมการจำแนกกลุ่มพบว่าทำนายได้ถูกต้องจำนวน 0 ตัวอย่าง คิดเป็นพยากรณ์ถูกต้อง 0.0 % ดังนั้นสมการจำแนกกลุ่มสามารถจำแนกได้ถูกต้องจำนวน 169 ตัวอย่าง คิดเป็น 84.5 %



ตารางที่ 4.3 ผลลัพธ์การจำแนกกลุ่มด้วยข้อมูลภาพที่ไม่มีการลดขนาดของวิธีซัพพอร์ตเวกเตอร์

ฟังก์ชัน เคอร์เนล	การแบ่งข้อมูล	กลุ่มจริง	การทำนายกลุ่ม		รวม	ความแม่นยำ
			1	2		
เชิงเส้น	Self-consistency	1	169 (100.0%)	0 (0%)	169	100.0%
		2	0 (0.0%)	31 (100.0%)	31	
	Split	1	48 (100.0%)	0 (0.0%)	48	93.3%
		2	4 (33.3%)	8 (66.7%)	12	
	Leave-one-out	1	169 (100.0%)	0 (0.0%)	169	96.0%
		2	8 (25.8%)	23 (74.2%)	31	
โพลีโนเมียล	Self-consistency	1	168 (100.0%)	1 (0.6%)	169	100.0%
		2	0 (0.0%)	31 (100.0%)	31	
	Split	1	48 (100.0%)	0 (0.0%)	48	80.0%
		2	12 (100.0%)	0 (0.0%)	12	
	Leave-one-out	1	169 (100.0%)	0 (0.0%)	169	84.5%
		2	31 (100.0%)	0 (0.0%)	31	

ตารางที่ 4.4 ผลลัพธ์การจำแนกกลุ่มด้วยข้อมูลภาพที่มีการลดขนาดของวิธีซัพพอร์ตเวกเตอร์

ฟังก์ชัน เคอร์เนล	การแบ่งข้อมูล	กลุ่มจริง	การทำนายกลุ่ม		รวม	ความแม่นยำ
			1	2		
เชิงเส้น	Self-consistency	1	168 (100.0%)	1 (0.6%)	169	100.0%
		2	0 (0.0%)	31 (100.0%)	31	
	Split	1	48 (100.0%)	0 (0.0%)	48	80.0%
		2	12 (100.0%)	0 (0.0%)	12	
	Leave-one-out	1	169 (100.0%)	0 (0.0%)	169	84.5%
		2	31 (100.0%)	0 (0.0%)	31	
โพลีโนเมียล	Self-consistency	1	168 (100.0%)	1 (0.6%)	169	100.0%
		2	0 (0.0%)	31 (100.0%)	31	
	Split	1	48 (100.0%)	0 (0.0%)	48	80.0%
		2	12 (100.0%)	0 (0.0%)	12	
	Leave-one-out	1	169 (100.0%)	0 (0.0%)	169	84.5%
		2	31 (0.0%)	0 (100.0%)	31	

#### 4.4 ผลการประเมินผลประสิทธิภาพ

งานวิจัยนี้มีการประเมินประสิทธิภาพการจำแนกกลุ่มคุณภาพในการตรวจสอบเส้นดำซึ่งเป็นข้อบกพร่องของกึ่งแข็งโดยใช้เทคนิคการสร้างโมเดลการจำแนกที่มีขั้นตอนและกระบวนการที่ต่างกัน ดังนั้นการวัดประสิทธิภาพนั้นสามารถพิจารณาได้จากค่าความถูกต้องจากการทำนายด้วยผลการทดสอบโมเดลและเปรียบเทียบผลของแต่ละโมเดลซึ่งสามารถคำนวณได้จากค่าในตารางเมทริกซ์ความสับสน (confusion Matrix) ที่มีลักษณะข้อมูลดังตารางที่ 3.2 ซึ่งเป็นลักษณะข้อมูลเดียวกับตารางผลลัพธ์การจำแนกกลุ่มตารางที่ 4.1, 4.2, 4.3 และ 4.4 แล้วนำค่าดังกล่าวมาคำนวณเพื่อประเมินประสิทธิภาพเป็นร้อยละ ประกอบด้วย ค่าความแม่นยำของโมเดล (accuracy) ดังสมการที่ 3.9 ค่าของโมเดลที่ทำนายได้ถูกต้อง (precision) ดังสมการที่ 3.10 ค่าจากการทำนายด้วยโมเดลที่ตรงกับความจริง (recall) ดังสมการที่ 3.11 และค่าเปรียบเทียบระหว่างค่า precision และค่า recall ของแต่ละกลุ่มเป้าหมาย (F-measure) การประเมินประสิทธิภาพของแต่ละโมเดลมีผลในตารางที่ 4.5 และ 4.6 ดังนี้

ตารางที่ 4.5 ผลการประเมินประสิทธิภาพการจำแนกกลุ่มคุณภาพด้วยวิธี discriminant analysis

โมเดล	การแบ่งข้อมูล	ไม่ลดขนาดข้อมูล						ลดขนาดข้อมูล					
		Accuracy	Precision		Recall		F-Measure	Accuracy	Precision		Recall		F-Measure
			Positive	Negative	Positive	Negative			Positive	Negative			
Discriminant Analysis	Self-consistency	0.97	0.96	0	1.00	0.18	0.98	0.93	0.92	0	1	0.31	0.96
	Split	0.93	0.93	0	1.00	0.40	0.96	0.90	0.85	0	1	0.57	0.92
	Cross-validated	0.96	0.95	0	1.00	0.23	0.97	0.93	0.92	0	1	0.31	0.96

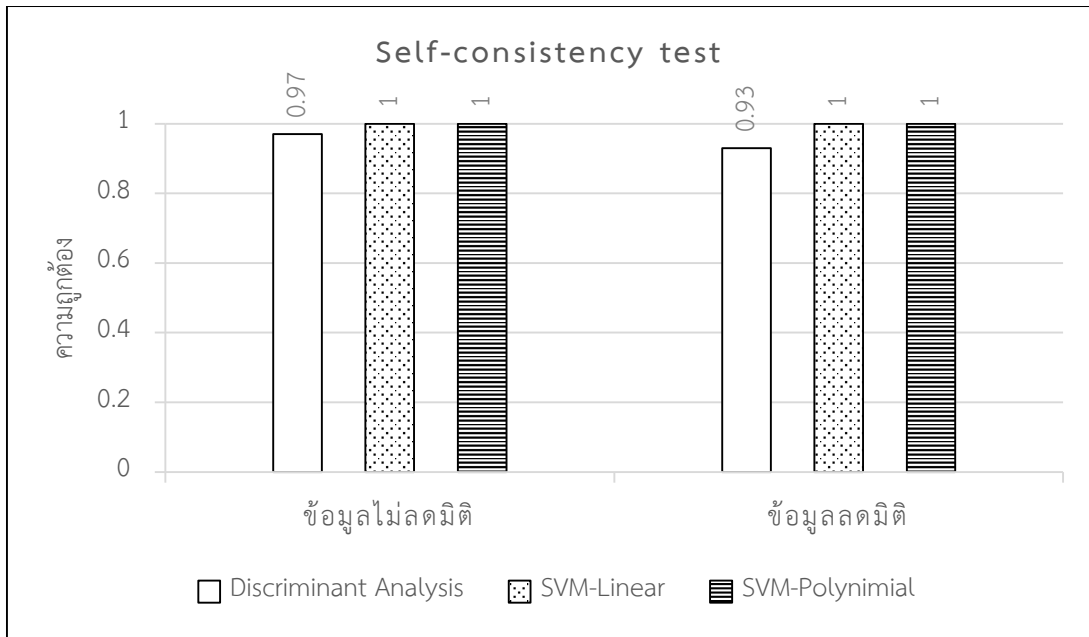
ตารางที่ 4.6 ผลการประเมินประสิทธิภาพการจำแนกกลุ่มคุณภาพวิธีซัพพอร์ตเวกเตอร์แมชชีน

โมเดล	การแบ่งข้อมูล	ไม่ลดขนาดข้อมูล						ลดขนาดข้อมูล					
		Accuracy	Precision		Recall		F-Measure	Accuracy	Precision		Recall		F-Measure
			Positive	Negative	Positive	Negative			Positive	Negative			
เชิงเส้น	Self-consistency	1.00	1.00	0	1.00	0	1.00	1.00	0.99	0	1.00	0.03	0.99
	Split	0.93	1.00	0.33	0.92	0	0.96	0.8	1.00	1.00	0.80	0	0.89
	Cross-validated	0.96	1.00	0.26	0.96	0	0.98	0.85	1.00	1.00	0.845	0	0.92
โพลีเมียล	Self-consistency	1	0.99	0	1.00	0.03	0.99	1.00	0.99	0	1.00	0.03	0.99
	Split	0.8	1.00	1.00	0.8	0	0.89	0.8	1.00	1.00	0.80	0	0.89
	Cross-validated	0.85	1.00	1.00	0.85	0	0.92	0.85	1.00	1.00	0.85	0	0.92

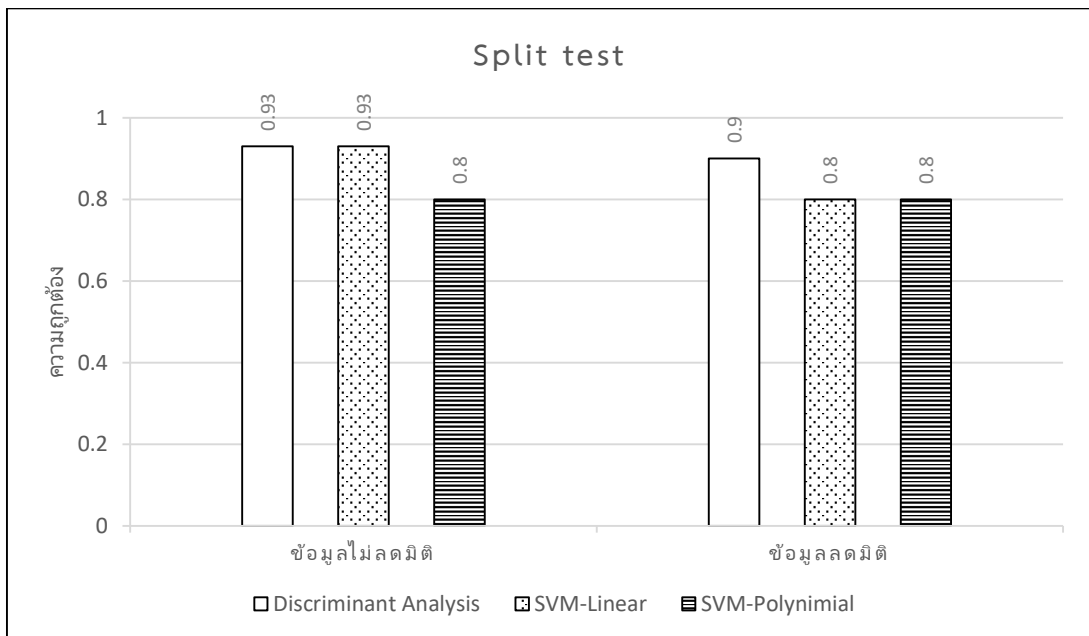
ผลคำนวณเพื่อประเมินประสิทธิภาพการจำแนกกลุ่มคุณภาพด้วยการแบ่งข้อมูลวิธี self-consistency test พบว่าเมื่อใช้ข้อมูลที่ไม่มีการลดขนาด ค่าความแม่นยำของโมเดลของวิธี discriminant analysis วิธีซัพพอร์ตเวกเตอร์แมชชีนฟังก์ชันเคอร์เนลแบบเชิงเส้น และแบบโพลีโนเมียล มีค่าเท่ากับร้อยละ 0.97 ,1.00 และ 1.00 เมื่อใช้ข้อมูลลดขนาดค่าความแม่นยำของโมเดล มีค่าเท่ากับร้อยละ 0.93, 1.00 และ 1.00 ตามลำดับ ผลการเปรียบเทียบค่าความแม่นยำดังรูปที่ 4.14(ก) แสดงให้เห็นว่าวิธีซัพพอร์ตเวกเตอร์แมชชีนฟังก์ชันเคอร์เนลแบบเชิงเส้น และโพลีโนเมียลทั้งกรณีที่ใช้ข้อมูลไม่ลดขนาดและลดขนาดมีความแม่นยำมากที่สุด

การแบ่งข้อมูลวิธี split test พบว่าเมื่อใช้ข้อมูลที่มีการลดขนาด ค่าความแม่นยำของโมเดลของวิธี discriminant analysis วิธีซัพพอร์ตเวกเตอร์แมชชีนฟังก์ชันเคอร์เนลแบบเชิงเส้น, และแบบโพลีโนเมียล มีค่าเท่ากับร้อยละ 0.93 ,0.93 และ 0.80 ตามลำดับ เมื่อใช้ข้อมูลลดขนาดค่าความแม่นยำของโมเดล มีค่าเท่ากับร้อยละ 0.90, 0.80 และ 0.80 ตามลำดับ ผลการเปรียบเทียบค่าความแม่นยำดังรูปที่ 4.14(ข) แสดงให้เห็นว่าวิธี discriminant analysis และวิธีซัพพอร์ตเวกเตอร์แมชชีนฟังก์ชันเคอร์เนลแบบเชิงเส้นกรณีที่ใช้ข้อมูลไม่ลดขนาดมีความแม่นยำมากที่สุด

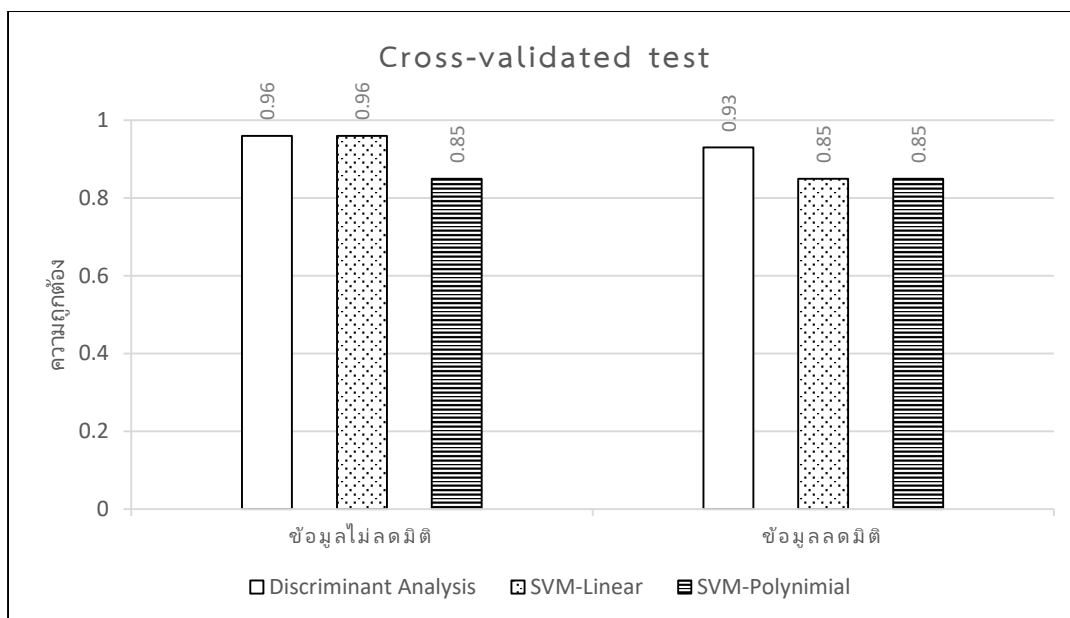
การแบ่งข้อมูลวิธี cross-validated test พบว่าเมื่อใช้ข้อมูลที่มีการลดขนาด ค่าความแม่นยำของโมเดลของวิธี discriminant analysis วิธีซัพพอร์ตเวกเตอร์แมชชีนฟังก์ชันเคอร์เนลแบบเชิงเส้น และแบบโพลีโนเมียล มีค่าเท่ากับร้อยละ 0.96 ,0.96 และ 0.85 ตามลำดับ เมื่อใช้ข้อมูลลดขนาดค่าความแม่นยำของโมเดล มีค่าเท่ากับร้อยละ 0.93, 0.85 และ 0.85 ตามลำดับ ผลการเปรียบเทียบค่าความแม่นยำดังรูปที่ 4.14(ค) แสดงให้เห็นว่าวิธี discriminant analysis และวิธีซัพพอร์ตเวกเตอร์แมชชีนฟังก์ชันเคอร์เนลแบบเชิงเส้นกรณีที่ใช้ข้อมูลไม่ลดขนาดมีความแม่นยำมากที่สุด



(ก) การแบ่งข้อมูลวิธี self-consistency test



(ข) การแบ่งข้อมูลวิธี split test



(ค) การแบ่งข้อมูลวิธี cross-validated test

รูปที่ 4.13 ผลเปรียบเทียบประสิทธิภาพด้านความถูกต้อง (accuracy)

ผลการคำนวณของค่าเปรียบเทียบระหว่างค่าของโมเดลที่ทำนายได้ถูกต้อง (precision) และค่าจากการทำนายด้วยโมเดลที่ตรงกับความจริง (recall) ของกลุ่มผ่านมาตรฐาน (F-measure) จากตารางที่ 4.5 และ 4.6 พบว่าการแบ่งข้อมูลวิธี self-consistency test วิธีซัพพอร์ตเวกเตอร์แมชชีนฟังก์ชันเคอร์เนลแบบเชิงเส้นกรณีข้อมูลที่ไม่มีลดขนาดจำแนกกลุ่มคุณภาพได้ดีที่สุด ใกล้เคียงกับฟังก์ชันเคอร์เนลแบบโพลีโนเมียลกรณีข้อมูลที่ไม่มีลดขนาด การแบ่งข้อมูลด้วยวิธี split test พบว่าวิธี discriminant analysis และวิธีซัพพอร์ตเวกเตอร์แมชชีนฟังก์ชันเคอร์เนลแบบเชิงเส้นกรณีข้อมูลที่ไม่มีลดขนาดจำแนกกลุ่มคุณภาพได้ดีที่สุด การแบ่งข้อมูลด้วยวิธี cross-validated test พบว่าวิธีซัพพอร์ตเวกเตอร์แมชชีนฟังก์ชันเคอร์เนลแบบเชิงเส้น และวิธี discriminant analysis กรณีข้อมูลที่ไม่มีลดขนาดจำแนกกลุ่มคุณภาพได้ดีที่สุด

การประเมินผลประสิทธิภาพของแต่ละวิธีในการจำแนกกลุ่มจำเป็นต้องเปรียบเทียบตามการแบ่งข้อมูลในการสร้างและทดสอบโมเดล เนื่องจากการแบ่งข้อมูลวิธีการเดียวกันผลประสิทธิภาพอาจจะต่างกันซึ่งขึ้นกับลักษณะของข้อมูลที่สุ่มได้ ทำให้วิธีการแบ่งข้อมูลจึงส่งผลกระทบต่อความแม่นยำของโมเดล จากการศึกษาวิจัยพบว่าโมเดลการจำแนกที่น่าเชื่อถือวิธีการในการแบ่งข้อมูลที่มีการสุ่มข้อมูลเพื่อให้โมเดลเรียนรู้ และวัดประสิทธิภาพมากที่สุด นั่นคือวิธี cross-validated test จากงานวิจัยนี้จะได้ว่าโมเดลการจำแนกกลุ่มวิธีซัพพอร์ตเวกเตอร์แมชชีนฟังก์ชันเคอร์เนลแบบเชิงเส้นโดยใช้ข้อมูลรูปแบบที่ไม่ลดขนาดมีความแม่นยำมากที่สุด แต่โมเดลการจำแนกที่นิยมใช้วัดผลเป็น

วิธีการในการแบ่งข้อมูล นั่นคือวิธี split test ซึ่งจะได้ว่าโมเดลการจำแนกกลุ่มวิธี discriminant analysis และวิธีซัพพอร์ตเวกเตอร์แมชชีนฟังก์ชันเคอร์เนลแบบเชิงเส้น โดยใช้ข้อมูลรูปแบบที่ไม่ลดขนาดมีความแม่นยำมากที่สุด นอกจากนี้อาจจะพิจารณารูปแบบของข้อมูลควบคู่กับวิธีในการจำแนก ซึ่งข้อมูลรูปแบบที่มีและไม่มีการลดขนาดจะแตกต่างกันในการใช้ระยะเวลาประมวลผล โดยข้อมูลที่ไม่มีการลดขนาดจะใช้เวลาในการประมวลผลมากกว่าข้อมูลที่มีการลดขนาด

ในการพัฒนาเทคนิคการตรวจสอบข้อบกพร่องของกึ่งโดยประยุกต์ใช้เทคนิคประมวลผลภาพ และอัลกอริทึมการจำแนกกลุ่มซึ่งมีผลการดำเนินงานตามขั้นตอนและเปรียบเทียบประสิทธิภาพในการจำแนกคุณภาพข้างต้น ดังนั้นจึงสามารถสรุปและอภิปรายผลการดำเนินงานของวิจัยนี้โดยมีรายละเอียดในบทที่ 5

## บทที่ 5

### บทสรุปและข้อเสนอแนะ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาเทคนิคการตรวจสอบเส้นด้ายซึ่งเป็นข้อบกพร่องของกุ้งแช่แข็งโดยประยุกต์ใช้เทคนิคประมวลผลภาพ และประเมินประสิทธิภาพในการจำแนกให้มีความถูกต้องในการตรวจสอบอย่างน้อย 95% เมื่อเทียบกับลักษณะตามมาตรฐานของผลิตภัณฑ์นำเข้าของประเทศคู่ค้าที่แบ่งระดับคุณภาพไว้ 2 ระดับ ประกอบด้วย ระดับผ่าน และระดับไม่ผ่านเกณฑ์มาตรฐาน สามารถสรุปผลการดำเนินการและข้อเสนอแนะสำหรับการพัฒนาเทคนิคได้ดังนี้

#### 5.1 สรุปผลการดำเนินงานวิจัย

จากวัตถุประสงค์ที่ได้กล่าวมาแล้วข้างต้นจึงได้พัฒนาเทคนิคการตรวจสอบข้อบกพร่องของกุ้งโดยใช้ตัวอย่างกุ้งขาว (*Litopenaeus vanamei*) ซึ่งเป็นสายพันธุ์กุ้งที่มีสัดส่วนของปริมาณการใช้เป็นวัตถุดิบในผลิตภัณฑ์กุ้งสดแช่แข็งมากที่สุด เก็บภาพกุ้งด้วยกล้องบันทึกภาพดิจิทัลในแต่ละระดับคุณภาพตามเกณฑ์มาตรฐานการนำเข้ากุ้งแช่แข็งของประเทศสหรัฐอเมริกา ซึ่งถูกแบ่งไว้ 2 ระดับ จำนวนทั้งหมด 200 ตัวอย่าง ประกอบด้วยกุ้งที่ผ่านเกณฑ์มาตรฐานจำนวน 31 ตัวอย่าง และกุ้งไม่ผ่านมาตรฐานจำนวน 169 ตัวอย่าง จึงกำหนดกลุ่มคุณภาพสำหรับจำแนกจำนวน 2 กลุ่มคือ กลุ่มที่ 1 กลุ่มผ่านเกณฑ์มาตรฐาน และกลุ่มที่ 2 กลุ่มไม่ผ่านเกณฑ์มาตรฐาน โดยการดึงลักษณะเด่นของข้อบกพร่องด้วยเทคนิคประมวลผลภาพซึ่งมี 3 ขั้นตอน ได้แก่ การประมวลผลภาพเบื้องต้น การแยกบริเวณภาพ และการแยกคุณลักษณะ โดยคุณลักษณะเฉพาะที่สามารถสกัดออกมาได้จากภาพและผ่านการทดสอบทางสถิติเพื่อนำมาเป็นพารามิเตอร์สำหรับการจำแนกซึ่งแบ่งออกเป็น 2 ประเภท ดังนี้

- 1) คุณลักษณะทางรูปร่าง ประกอบด้วย พื้นที่ และความยาว
- 2) คุณลักษณะจากฮิสโตแกรม ประกอบด้วย ค่าเฉลี่ย ค่าความสว่างต่ำสุด ค่าเบี่ยงเบนมาตรฐาน และเปอร์เซ็นต์ไทล์ที่ 25 และ 50

ในงานวิจัยนี้ได้นำเทคนิค Principal Component Analysis (PCA) ซึ่งเป็นการลดขนาดข้อมูลมาใช้ในการเตรียมข้อมูลให้ข้อมูลมีจำนวนตัวแปรน้อยลง โดยยังคงลักษณะเด่นและสูญเสียความถูกต้องของผลลัพธ์น้อยที่สุด นั่นคือการรวมคุณลักษณะมีความสัมพันธ์กันมากไว้เป็นปัจจัย



เดียวกันทำให้ได้ข้อมูลชุดใหม่ด้วยการจัดกลุ่มปัจจัย (component) ร่วมที่เหมาะสม สามารถจัดกลุ่มพารามิเตอร์จากจำนวน 7 ปัจจัยให้เหลือ 2 ปัจจัยด้วยสมการที่ 4.1 และ 4.2 ซึ่งทั้ง 2 ปัจจัยดังกล่าวประกอบด้วยพารามิเตอร์ดังนี้

- ปัจจัยที่ 1 ประกอบด้วยตัวแปร 5 พารามิเตอร์คือ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน พื้นที่ ค่าความสว่างที่น้อยที่สุด และความยาว

- ปัจจัยที่ 2 ประกอบด้วยตัวแปร 2 พารามิเตอร์คือ เพอร์เซ็นต์ไทล์ที่ 25 และ เพอร์เซ็นต์ไทล์ที่ 50

จากนั้นนำข้อมูลทั้ง 2 กรณี คือ กรณีที่มีการลดขนาดข้อมูล และไม่มีการลดขนาดข้อมูลมาแบ่งข้อมูลสำหรับการวัดประสิทธิภาพของเทคนิคการจำแนกเพื่อการสอนและทดสอบโมเดล โดยใช้กลุ่มข้อมูลสำหรับการสอนเพื่อสร้างโมเดล และใช้กลุ่มข้อมูลสำหรับการทดสอบความแม่นยำของการจำแนกคุณภาพ 3 วิธีด้วยกัน คือ วิธี self-consistency test วิธี split test และวิธี cross-validation test

เมื่อนำข้อมูลในแต่ละกรณีข้างต้นมาใช้ในการจำแนกกลุ่มระดับคุณภาพตามเกณฑ์มาตรฐานด้วย 2 วิธี คือ วิธีที่ 1 คือ วิธี discriminant analysis โดยการวิเคราะห์การจำแนกกลุ่มวิธีนี้จะสร้างสมการตามวิธี Fisher's linear discrimination function จากสมการของข้อมูลที่ไม่มีการลดขนาดดังสมการที่ 4.3 และ 4.4 ประกอบด้วยพารามิเตอร์คือ พื้นที่ ความยาว ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และ เพอร์เซ็นต์ไทล์ที่ 50 ส่วนสมการของข้อมูลที่มีการลดขนาดดังสมการที่ 4.5 และ 4.6 ประกอบด้วยพารามิเตอร์คือ ปัจจัยที่ 1 และปัจจัยที่ 2 ที่ได้จากสมการค่าคะแนนปัจจัยจากสมการการลดข้อมูล และวิธีที่ 2 คือ วิธีซัพพอร์ตเวกเตอร์แมชชีนโดยการวิเคราะห์การจำแนกกลุ่มวิธีนี้เป็นการสร้างเส้นแบ่งแยกกลุ่มข้อมูลด้วยการหาสัมประสิทธิ์ของสมการโมเดล จากข้อมูลภาพที่ไม่ผ่านกระบวนการลดขนาดมีพารามิเตอร์สำหรับสร้างโมเดล คือ พื้นที่ ความยาว ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าต่ำสุด เพอร์เซ็นต์ไทล์ที่ 25 และ 50 ส่วนข้อมูลภาพที่ผ่านกระบวนการลดขนาด มีพารามิเตอร์สำหรับสร้างโมเดล 2 พารามิเตอร์ที่ได้จากสมการค่าคะแนนปัจจัย

ผลประสิทธิภาพของแต่ละวิธีในการจำแนกกลุ่มเปรียบเทียบความแม่นยำตามการแบ่งข้อมูลในการสร้างและทดสอบโมเดล เนื่องจากการแบ่งข้อมูลที่แตกต่างกันข้อมูลที่ได้ในการสร้างและทดสอบก็อาจจะแตกต่างกัน ผลประสิทธิภาพอาจจะต่างกันเช่นเดียวกัน ซึ่งขึ้นกับลักษณะของข้อมูลที่สุ่มได้ ทำให้วิธีการแบ่งข้อมูลจึงส่งผลต่อความแม่นยำของโมเดล โดยการแบ่งกลุ่มข้อมูลวิธีการแบ่งข้อมูลแบบ self-consistency test วิธีการจำแนกที่มีความแม่นยำของโมเดลมากที่สุด คือ วิธีซัพพอร์ตเวกเตอร์แมชชีนฟังก์ชันเคอร์เนลแบบเชิงเส้น และโพลิโนเมียลกรณีข้อมูลที่ไม่มีการลดขนาดมีความถูกต้องที่ 100.0% วิธีการแบ่งข้อมูลแบบ split test วิธีการจำแนกที่มีความแม่นยำของโมเดลมากที่สุด คือ วิธี discriminant analysis กรณีข้อมูลที่ไม่มีการลดขนาดมีความถูกต้องที่

93.3% ซึ่งใกล้เคียงกับวิธีซัพพอร์ตเวกเตอร์แมชชีนฟังก์ชันเคอร์เนลแบบเชิงเส้น และโพลิโนเมียลกรณีข้อมูลที่ไม่มีการลดขนาดที่มีความถูกต้องที่ 93.0 % และวิธีการแบ่งข้อมูลแบบ cross-validated test วิธีการจำแนกที่มีความแม่นยำของโมเดลมากที่สุด คือ วิธีซัพพอร์ตเวกเตอร์แมชชีนฟังก์ชันเคอร์เนลแบบเชิงเส้นกรณีข้อมูลที่ไม่มีการลดขนาดมีความถูกต้องที่ 96.5% ซึ่งวิธีการแบ่งข้อมูลแบบ cross-validated test นี้เป็นวิธีแบ่งข้อมูลสำหรับการทดสอบที่น่าเชื่อถือ เนื่องจากมีการสุ่มข้อมูลในการทดสอบหลายครั้ง แต่การแบ่งข้อมูลที่นิยมใช้วัดผลของโมเดล นั่นคือวิธี split test ซึ่งจะเป็นแบ่งข้อมูลเพียงครั้งเดียว โดยถือว่าหากโมเดลที่สร้างมีประสิทธิภาพจะสามารถทำนายผลได้ถูกต้องแม่นยำทุกกลุ่มตัวอย่าง จึงนิยมใช้ในงานวิจัยมากที่สุด

ดังนั้นจากงานวิจัยนี้จะได้ว่าโมเดลการจำแนกกลุ่มวิธีซัพพอร์ตเวกเตอร์แมชชีนฟังก์ชันเคอร์เนลแบบเชิงเส้นเป็นวิธีการที่ดีที่สุด เนื่องจากสามารถทำนายผลได้มีความแม่นยำมากในทุกรูปแบบการแบ่งข้อมูล นอกจากนี้การลดขนาดของข้อมูลให้มีจำนวนลดลงทำให้ความถูกต้องของทำนายลดลง แต่ระยะเวลาในการวิเคราะห์แตกต่างกันเพียงเล็กน้อย จึงไม่จำเป็นต้องใช้ข้อมูลที่มีการลดขนาดในการสร้างโมเดล

## 5.2 ข้อเสนอแนะในการดำเนินงานวิจัย

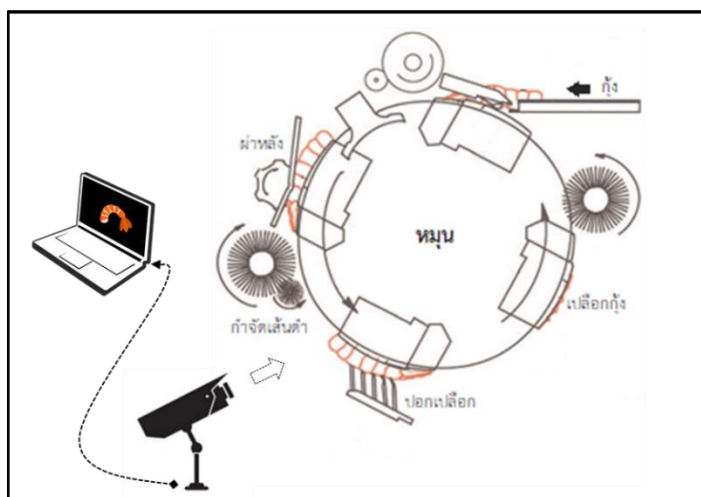
จากการวิจัยข้างต้นทำให้ทราบถึงคุณลักษณะที่เหมาะสมสำหรับการจำแนกลักษณะบกพร่องของการกำจัดเส้นดำไม่ถูกต้องออกจากผลิตภัณฑ์ที่ตรงตามเกณฑ์มาตรฐาน และขั้นตอนวิธีการในเก็บข้อมูลจนได้พารามิเตอร์ที่เหมาะสม รวมทั้งวิธีการและเทคนิคในการคัดแยกที่ให้ความถูกต้องมากที่สุดจากการทดลอง ซึ่งวิธีการและเทคนิคที่ได้สามารถนำไปสร้างระบบอัตโนมัติและประยุกต์ใช้จริงในอนาคตได้ จากการวิจัยมีข้อเสนอแนะในการดำเนินงานวิจัยและแนวทางในการพัฒนาเพื่อนำผลการศึกษาไปประยุกต์ใช้ดังนี้

1) การประยุกต์ใช้เทคนิคที่พัฒนาขึ้นในการตรวจสอบลักษณะบกพร่องของกึ่งจำเป็นต้องควบคุมปัจจัยต่างๆที่เกี่ยวข้อง คือ การกำหนดค่าเริ่มต้นของกล้องถ่ายภาพ การวางตำแหน่งของกล้องถ่ายภาพ แสงสว่าง ซึ่งต้องคำนึงถึงชนิด จำนวน และการตำแหน่งการติดตั้งของหลอดไฟที่ใช้เป็นแหล่งของแสงสว่าง อีกทั้งควรใช้กล่องควบคุมแสงเพื่อควบคุมแสงสว่างให้มีสภาวะคงที่ และทำให้เทคนิคดังกล่าวสามารถตรวจสอบระดับลักษณะบกพร่องของกึ่งได้อย่างถูกต้องและรวดเร็ว

2) การเก็บภาพกึ่งเข้าระบบจำแนกต้องป้อนครั้งละ 1 ตัว และจัดวางกึ่งในลักษณะที่สามารถมองเห็นบริเวณหลังที่ถูกกำจัดเส้นดำได้ก่อนการถ่ายภาพ หากจะเก็บภาพบนสายพานจำเป็นต้องมีพนักงานจัดวางกึ่งก่อนการเก็บภาพ

3) เทคนิคที่พัฒนาสามารถปรับเปลี่ยนคุณลักษณะ และวิธีการจำแนกกลุ่มให้เหมาะสมตามเกณฑ์มาตรฐาน รวมทั้งสายพันธุ์หรือขนาดกึ่งที่แตกต่างกัน

4) สามารถนำมาประยุกต์ใช้กระบวนการและเทคนิคที่พัฒนาขึ้นกับเทคโนโลยีอัตโนมัติในกระบวนการผลิตปัจจุบัน เช่น เครื่องทำความสะอาด เครื่องคัดขนาด เครื่องปอกเปลือก ผ่าหลัง และกำจัดเส้นดำ เป็นต้น เพื่อลดต้นทุนการผลิต ควบคุมมาตรฐาน และตรวจสอบคุณภาพของผลิตภัณฑ์ และนอกจากนี้ยังมีผู้วิจัยจากหลายสถาบันได้สร้างระบบอัตโนมัติในการกระบวนการผลิตกึ่ง [71], [72] ดังนั้นจึงมีแนวคิดนำมาประยุกต์ใช้กับเครื่องจักรอัตโนมัติที่ทำการปอกเปลือก ผ่าหลัง และกำจัดเส้นดำกึ่ง ซึ่งกึ่งจะผ่านเครื่องมือภายในเครื่องจักร จากรูปที่ 5.1 เมื่อกึ่งเข้าเครื่องจักรจะถูกผ่าหลังแล้วใช้แปรงขัดเพื่อกำจัดเส้นดำและดึงตัวกึ่งออกจากเปลือกด้วยเครื่องมือเกี่ยวเปลือกออก การทำงานดังกล่าวสามารถนำเทคนิคประมวลผลภาพประยุกต์ใช้ได้โดยติดตั้งกล้อง ณ ตำแหน่งหลังจากขั้นตอนกำจัดเส้นดำแล้ว เพราะขั้นตอนต่างๆจะเสร็จสิ้นภายในเครื่องจักร และมีโลหะหนีบตัวกึ่งไว้ในลักษณะที่สามารถถ่ายรูปให้เห็นส่วนหลังที่มีการกำจัดเส้นดำได้อย่างชัดเจน



รูปที่ 5.1 การประยุกต์เทคนิคประมวลผลภาพกับเครื่องจักรอัตโนมัติ

5) สำหรับการปรับค่าขีดแบ่ง (thresholding value) ให้เหมาะสมตามฮิสโตแกรมของภาพในขั้นตอนการแบ่งส่วนภาพ (segmentation) จากการเก็บข้อมูลภาพตัวอย่างจำนวน 200 ภาพ พบว่าปรับค่าขีดแบ่งมีค่าเฉลี่ย ส่วนตัวกึ่งให้ค่าอยู่ระหว่าง 79 ถึง 154 และส่วนเส้นดำปรับค่าขีดแบ่ง ให้อยู่ระหว่าง 83 ถึง 255

งานวิจัยนี้สามารถนำวิธีการและเทคนิคไปประยุกต์และเป็นแนวทางเพื่อการพัฒนาตรวจสอบและจำแนกลักษณะคุณภาพทางการเกษตรอื่นๆได้ นอกเหนือจากผลิตภัณฑ์ของกึ่ง เช่น

ตำหนิบนผิวของผลไม้ สีของผลิตภัณฑ์ที่ใช้ในการคัดแยกคุณภาพ เป็นต้น โดยใช้พารามิเตอร์และเทคนิคการจำแนกที่เหมาะสม

### บรรณานุกรม

- [1] MGR Online, “สถาบันอาหาร คาดตัวเลขส่งออกอาหารปี 60 ทะลุ 1 ล้านล้านบาท”, *อัจฉริยะเพื่ออุตสาหกรรมอาหาร (Food Intelligence Center)*, 16-ธค-2016.
- [2] หาญหิรัญสมชาย, “อุตสาหกรรมอาหารไทย ยังครองแชมป์ส่งออกอันดับ 1 ของโลก”, 18-ตค-2016.
- [3] เอนกพร เกษมสุข, “กุ้งสด แช่เย็น แช่แข็ง และแปรรูป (ไม่รวมกุ้งกระป๋อง)”, กรมส่งเสริมการค้าระหว่างประเทศ, สำนักพัฒนาการค้าและธุรกิจการเกษตรและอุตสาหกรรม, 2560.
- [4] สำนักวิจัยเศรษฐกิจการเกษตร, “สถานการณ์สินค้าเกษตรที่สำคัญและแนวโน้ม ปี 2560”, กระทรวงเกษตรและสหกรณ์, สำนักงานเศรษฐกิจการเกษตร, 2559.
- [5] นิรชา วงษ์จินดา, “กระบวนการผลิตกุ้งแช่เยือกแข็ง”. กรมประมง, 2556.
- [6] NOAA Fisheries, “Inspector’s Instructions for Grading Fresh and Frozen Shrimp”, ใน *NOAA Handbook, Seafood Inspection Program: NOAA U.S Department of Commerce*, 2014, น. 118–126.
- [7] NOAA Fisheries, “Inspector’s Guidelines for Grading Fresh and Frozen Shrimp”, ใน *The standards for grades, Seafood Inspection Program: NOAA U.S Department of Commerce*, 1993, น. 118–126.
- [8] L. Dah-Jye, G. Xiong, R. M. Lane และ D. Zhang, *An efficient shape analysis method for shrimp quality evaluation*. 2012.
- [9] Crushed Velvet, *Shrimp Size Chart*. 2016.
- [10] กรมเจรจาการค้าระหว่างประเทศ, “อุตสาหกรรมอาหารแปรรูป”, 09-มค-2016. [ออนไลน์]. Available at: [www.thaifita.com/thaifita/Portals/0/File/ascn\\_foosindus.doc](http://www.thaifita.com/thaifita/Portals/0/File/ascn_foosindus.doc).
- [11] ปพนวีณ์ สุทธิประสิทธิ์, “เทคนิคการตรวจสอบและการประเมินคุณภาพอาหารด้วยวิธีการทางประสาทสัมผัส”. มหาวิทยาลัยสุโขทัยธรรมาธิราช, 2557.
- [12] M. A. Somnuk Aujirapongpan และ Pongpan Compan, “A Study of Labor Demand on Frozen Seafood Entrepreneurs in Songkhla Province”, *WMS J. Manag. Walailak Univ.*, ปี 1, ฉบับที่ No.2 (Sep – Dec 2012), น. 1–8, 2012.
- [13] นุชรี ธรรมโชติ, “การพัฒนาระบบการประมวลผลภาพสำหรับกระบวนการคัดแยกขนาดและสายพันธุ์ของหมึกกล้วยแปรรูป”, มหาวิทยาลัยสงขลานครินทร์, 2557.
- [14] นิดดา มะสุณี, “การพัฒนาระบบการประมวลผลภาพสำหรับกระบวนการตรวจสอบคุณภาพของหมึกกล้วยแปรรูป”, มหาวิทยาลัยสงขลานครินทร์, 2557.

- [15] ศูนย์วิจัยกสิกรไทย, “ส่องทางรอด ส่งออกอาหารทะเลแปรรูปไทย”. ธนาคารกสิกรไทย, เมษายน-2017.
- [16] MM Thailand, “อุตสาหกรรม 4.0 จุดเปลี่ยนประเทศไทย”, *Modern Manufacturing*, 30-มีค-2017.
- [17] “เกษตร 4.0 เครื่องจักรผลิตกุ้งนวัตกรรมลดต้นทุนแรงงาน”, *พลังเกษตร*, 09-ธค-2016.
- [18] S. Hosseinpour, S. Rafiee, M. Aghbashlo และ S. S. Mohtasebi, “A novel image processing approach for in-line monitoring of visual texture during shrimp drying”, *J. Food Eng.*, ปี 143, ฉบับที่ Supplement C, น. 154–166, ธ.ค. 2014.
- [19] J.-H. Qu, J.-H. Cheng, D.-W. Sun, H. Pu, Q.-J. Wang และ J. Ma, “Discrimination of shelled shrimp (*Metapenaeus ensis*) among fresh, frozen-thawed and cold-stored by hyperspectral imaging technique”, *LWT - Food Sci. Technol.*, ปี 62, ฉบับที่ 1, Part 1, น. 202–209, มิ.ย. 2015.
- [20] D. Nagalakshmi, S. Jyothi และ P. D Scholar, *White Spot Syndrome Virus Detection in Shrimp Images using Image Segmentation Techniques*, ปี 3. 2013.
- [21] P. Pan, J. Li, G. Lv, H. Yang, S. Zhu และ J. Lou, “Prediction of shelled shrimp weight by machine vision”, *J. Zhejiang Univ. Sci. B*, ปี 10, ฉบับที่ 8, น. 589–594, ส.ค. 2009.
- [22] Poonpat Poonnoy และ Thinnapop Chum-in, “Estimation of Sushi Shrimp Weight Using Image Analysis Technique and Non-linear Regression Models”, Food Engineering Program, Faculty of Engineering and Agro-Industry, Maejo University,.
- [23] Murat O. Balaban, Ymir Bergmann และ Sencer Yeralan, “Determination of Count and Uniformity Ratio of Shrimp by Machine Vision”, *J. Aquat. Food Prod. Technol.*, ปี 3, ฉบับที่ 3, น. 43–58, 1994.
- [24] Poonnoy, P., Yodkeaw, P., Sriwai, A, Umongkol, P. และ Intamoon, S., “Classification of Boiled Shrimp’s Shape Using Image Analysis and Artificial Neural Network Model”, *J. Food Process Eng.*, ปี 37, ฉบับที่ 3, น. 257–263, มิ.ย. 2014.
- [25] D. Zhang, K. D. Lillywhite, D.-J. Lee และ B. J. Tippetts, “Automatic Shrimp Shape Grading Using Evolution Constructed Features”, *Comput Electron Agric*, ปี 100, น. 116–122, ม.ค. 2014.

- [26] Pairoj Klypetch, “A METHOD TO MEASURE THE COLOR OF THE PRINTED BANKNOTES USING DIGITAL IMAGE PROCESSING TO OPTIMIZE PRINT QUALITY”, Chulalongkorn University, 2013.
- [27] สำนักงานคณะกรรมการวิจัยแห่งชาติ, “การเลี้ยงกุ้งขาวแวนนาไม”, *Thailand Shrimp*, 20-มีย-2016. [ออนไลน์]. Available at: [http://www.thailandshrimp.org/agriculture\\_vannamei.html](http://www.thailandshrimp.org/agriculture_vannamei.html).
- [28] “กุ้ง - วิกีพีเดีย”. [ออนไลน์]. Available at : <https://th.wikipedia.org/wiki/%E0%B8%81%E0%B8%B8%E0%B9%89%E0%B8%87>. [สืบค้น: 24-ตค-2015].
- [29] ประจวบ หล้าอุบล, ข้อมูลพื้นฐานของกุ้ง. คณะประมง มหาวิทยาลัยเกษตรศาสตร์: สถาบันวิจัยการเพาะเลี้ยงสัตว์น้ำชายฝั่งสงขลา.
- [30] ประโรต ไทยเครือวัลย์, “บทที่ 12 กุ้งและแมลงสาบ”, ใน *สัตววิทยา (Zoology)*, สถาบันราชภัฏสงขลา, 2540.
- [31] “Shrimp”, *Dickinson College Department of Environmental Studies*, 17-กค-2016. [ออนไลน์]. Available at: <http://www2.dickinson.edu/departments/envst/luceweb/pages/hypoxia-gurbiszc/effects.htm>.
- [32] กมลศิริ พันธนิยะ, “กุ้งขาวลิโทพีเนียส แวนนาไม ((Pacific white shrimp)”, *Shrimp Center*, 01-ตค-2016. [ออนไลน์]. Available at: <http://www.shrimpcenter.com/t-shrimp051.html>.
- [33] “รายงานประจำปี 2558 บริษัท ห้างเย็นโซติวัฒน์หาดใหญ่ จำกัด(มหาชน)”, สงขลา, บริษัท ห้างเย็นโซติวัฒน์หาดใหญ่ จำกัด(มหาชน), 2558.
- [34] “กุ้ง ”, *Food Network Solution*, 25-มี ค - 2016. [ออนไลน์]. Available at: <http://www.foodnetworksolution.com/processing/category/100/%E0%B8%81%E0%B8%B8%E0%B9%89%E0%B8%87>.
- [35] “Product Gallery”, *Thai Frozen Foods Association (TFFA)*, 25-มีค-2016. [ออนไลน์]. Available at: [http://www.thai-frozen.or.th/product\\_gallery\\_shrimp.php](http://www.thai-frozen.or.th/product_gallery_shrimp.php).
- [36] “ความรู้เบื้องต้นเกี่ยวกับประมวลผลภาพ”, *NEXTSOFTWAREHOUSE*, 20-สค-2016. [ออนไลน์]. Available at: [http://www.thaifrozen.or.th/product\\_gallery\\_shrimp.php](http://www.thaifrozen.or.th/product_gallery_shrimp.php).
- [37] รังสฤษฎ์ คงด้วง, “ระบบอัตโนมัติสำหรับการประเมินปริมาณการใช้สีในกระบวนการผลิตผลิตภัณฑ์ไม่ย่างพารา”, มหาวิทยาลัยสงขลานครินทร์, 2558.

- [38] มนูญศักดิ์ วรรณประพันธ์และเอกพันธ์ อินดี, “โปรแกรมตรวจจับโมเดลส่วนท้ายรถยนต์แบบอัตโนมัติ”, มหาวิทยาลัยสยาม, 2556.
- [39] Kirkpong Kiatpanichagij, “Machine Vision”, 06-พย-2016. [ออนไลน์]. Available at: [http://www.geocities.ws/k\\_kirkpong/chapter01.pdf](http://www.geocities.ws/k_kirkpong/chapter01.pdf).
- [40] O. Marques, *Practical image and video processing using MATLAB*, ปี 2011. Hoboken, NJ: Wiley-IEEE Press.
- [41] Kirkpong Kiatpanichagij, “ฮิสโทแกรม”, จุฬาลงกรณ์มหาวิทยาลัย, 06-พย-2016. [ออนไลน์]. Available at: <http://pioneer.netserv.chula.ac.th/~kchawan/histogram/his.html>.
- [42] Darlene Hildebrandt, “How to Read and Use Histograms”, *Digital Photography School*, 17-กค-2016. [ออนไลน์]. Available at: <https://digital-photography-school.com/how-to-read-and-use-histograms/>.
- [43] เกริกพงษ์ เกียรติพานิชกิจ, “Introduction to Machine Vision”, น. 1-17.
- [44] ชินกร น้อยคำยาง, “การเปรียบเทียบผลการทดสอบสถิติพาราเมตริกและนอนพาราเมตริกในการทดสอบความแตกต่างระหว่างประชากรสองกลุ่มที่อิสระจากกัน”, มหาวิทยาลัยศรีนครินทรวิโรฒ, 2552.
- [45] “การทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ยของประชากร”, มหาวิทยาลัยขอนแก่น. [ออนไลน์]. Available at: [https://webcache.googleusercontent.com/search?q=cache:r3RU0JSK4foJ:https://home.kku.ac.th/sompo\\_pu/spweb/stat/02.doc+&cd=1&hl=th&ct=clnk&gl=th](https://webcache.googleusercontent.com/search?q=cache:r3RU0JSK4foJ:https://home.kku.ac.th/sompo_pu/spweb/stat/02.doc+&cd=1&hl=th&ct=clnk&gl=th)
- [46] สายชล สันสมบุรณ์ทอง, *การวิเคราะห์ตัวแปรหลายตัวโดยใช้ SPSS และ MINITAB*, 1 พิมพ์ครั้งที่. บริษัทจามจรีโปรดักส์ จำกัด: ศูนย์หนังสือจุฬาลงกรณ์มหาวิทยาลัย, 2559.
- [47] ปกรณ์ แก้วตระกูลพงษ์, *การประยุกต์อัลกอริทึมแมชชีนวิชันโดยใช้ NI Vision*, ปี 1. จุฬาลงกรณ์มหาวิทยาลัย, 2559.
- [48] “SVM - Support Vector Machines Optimum Separation Hyperplane”, *SVM - Support Vector Machines*, 16-กย-2017. [ออนไลน์]. Available at: [http://www.support-vector-machines.org/SVM\\_osh.html](http://www.support-vector-machines.org/SVM_osh.html).
- [49] Chanya Chuenarrom, “ImageJ 1: โปรแกรม ImageJ คืออะไร”, *ComDigest*, 15-ธค-2016. [ออนไลน์]. Available at: <http://comdigest.blogspot.com/2015/10/imagej.html>.



- [50] มาโนช ห้วยหงษ์ทอง, วชิรรัตน์ กุลทีและสรารัตน์ แก้วมานพ, “คู่มือแนะนำ WEKA”. มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี.
- [51] เกษทิพย์ ศิริชัยศิลป์, “โปรแกรม Spss คืออะไร”, 20-กค-2016. [ออนไลน์]. Available at: <https://www.gotoknow.org/posts/463011.%20%E0%B8%A7%E0%B8%B1%E0%B8%99%E0%B8%97%E0%B8%B5%E0%B9%88%E0%B8%AA%E0%B8%B7%E0%B8%9A%E0%B8%84%E0%B9%89%E0%B8%99%207%20%E0%B8%81%E0%B8%A3%E0%B8%81%E0%B8%8E%E0%B8%B2%E0%B8%84%E0%B8%A1%202557>.
- [52] H. Eerens, D. Haesen, F. Rembold, F. Urbano, C. Tote และ L. Bydekerke, “Image time series processing for agriculture monitoring”, *Environ. Model. Softw.*, ปี 53, ฉบับที่ Supplement C, น. 154–162, มี.ค. 2014.
- [53] D. S. Jayas, J. Paliwal และ N. S. Visen, “Review Paper (AE—Automation and Emerging Technologies): Multi-layer Neural Networks for Image Analysis of Agricultural Products”, *J. Agric. Eng. Res.*, ปี 77, ฉบับที่ 2, น. 119–128, ต.ค. 2000.
- [54] C.-J. Du และ D.-W. Sun, “Recent developments in the applications of image processing techniques for food quality evaluation”, *Trends Food Sci. Technol.*, ปี 15, ฉบับที่ 5, น. 230–249, พ.ค. 2004.
- [55] ประสิทธิ์ นครราช, จุริรัตน์ อ้วนศรีเมืองและนภาพรณัฏ มัธน์ง, “การศึกษาการตรวจสอบคุณภาพทางกายภาพของข้าวโดยการประมวลผลภาพดิจิทัลด้วยวิธีการ”, ใน *The 5th UBU Conference Proceeding*, มหาวิทยาลัยอุบลราชธานี, 2554, น. 194–202.
- [56] P. Moallem, A. Serajoddin และ H. Pourghassem, “Computer vision-based apple grading for golden delicious apples based on surface features”, *Inf. Process. Agric.*, ปี 4, ฉบับที่ 1, น. 33–40, มี.ค. 2017.
- [57] G. Jahns, H. Møller Nielsen และ W. Paul, “Measuring image analysis attributes and modelling fuzzy consumer aspects for tomato quality grading”, *Comput. Electron. Agric.*, ปี 31, ฉบับที่ 1, น. 17–29, มี.ค. 2001.
- [58] P. P. Ling และ V. N. Ruzhitsky, “Machine Vision Techniques for Measuring the Canopy of Tomato Seedling”, *J. Agric. Eng. Res.*, ปี 65, ฉบับที่ 2, น. 85–95, ต.ค. 1996.

- [59] J. Li, X. Rao และ Y. Ying, “Detection of common defects on oranges using hyperspectral reflectance imaging”, *Comput. Electron. Agric.*, ปี 78, ฉบับที่ 1, น. 38–48, ส.ค. 2011.
- [60] C. P. Diana Carolina และ N. T. Devis David, “Classification of oranges by maturity, using image processing techniques”, ใน *2014 III International Congress of Engineering Mechatronics and Automation (CIIMA)*, 2014, น. 1–5.
- [61] Y. Wu, D. Li, Z. Li และ W. Yang, “Fast processing of foreign fiber images by image blocking”, *Inf. Process. Agric.*, ปี 1, ฉบับที่ 1, น. 2–13, ส.ค. 2014.
- [62] N. Kondo, H. Murase, M. Monta, Y. Shibano และ K. Mohri, “Study on Quality Evaluation of Orange Fruit Using Image Processing”, *1st IFAC Workshop Control Appl. Post-Harvest Process. Technol. CAPPT95 Ostend Belg. 1-2 June 1995*, ปี 28, ฉบับที่ 6, น. 125–128, มิ.ย. 1995.
- [63] H. M. Velioğlu, İ. H. Boyacı และ Ş. Kurultay, “Determination of visual quality of tomato paste using computerized inspection system and artificial neural networks”, *Comput. Electron. Agric.*, ปี 77, ฉบับที่ 2, น. 147–154, ก.ค. 2011.
- [64] M. Weis และ R. Gerhards, *Detection of weeds using image processing and clustering*, ปี 69. 2009.
- [65] A. Ehsani Rad และ Y. H. Kumar, *Leaf recognition for plant classification using GLCM and PCA methods*, ปี 3. 2010.
- [66] D. Mas และคณะ, “Novel image processing approach to detect malaria”, *Opt. Commun.*, ปี 350, ฉบับที่ Supplement C, น. 13–18, ก.ย. 2015.
- [67] จักริน สุขสวัสดิ์ชน, อุรีรัฐ สุขสวัสดิ์ชน และ พงศ์พัฒน์ สิงห์ศรี, “การคัดเลือกปัจจัยเด่นและการเปรียบเทียบประสิทธิภาพเทคนิคการแยกประเภทสำหรับทำนายเพศของนกกระทาจากคุณลักษณะภายนอกของไข่นกกระทา”, นำเสนอที่ The Tenth National Conference on Computing and Information Technology, 2014, น. 515–521.
- [68] A. Mizushima และ R. Lu, “An image segmentation method for apple sorting and grading using support vector machine and Otsu’s method”, *Comput. Electron. Agric.*, ปี 94, ฉบับที่ Supplement C, น. 29–37, มิ.ย. 2013.
- [69] J.-H. C. Jia-Huan Qu และ H. P. Da-Wen Sun, “Discrimination of shelled shrimp (*Metapenaeus ensis*) among fresh, frozen-thawed and cold-stored by

- hyperspectral imaging technique”, *Food Science and Technology*, ปี 2015, ฉบับที่ 62, น. 202–209.
- [70] F. Pedreschi, J. León, D. Mery และ P. Moyano, “Development of a computer vision system to measure the color of potato chips”, *Food Research International*, ปี 2006, ฉบับที่ 39, น. 1092–1098.
- [71] ณัฐสิมา นาคบุตร, ธีัญญลักษณ์ สุทธนะและมณีรัตน์ เหลืองทรงชัย, “เครื่องผ่าหลังกุ้ง”, มหาวิทยาลัยเกษตรศาสตร์.
- [72] พูนพัฒน์ พูนน้อยและกิตติกร หาญตระกูล, “การพัฒนาเครื่องคัดขนาดกุ้งผ่าหลังแบบผล็ืออัตโนมัติระบบคอมพิวเตอร์วิสัยทัศน์”, มหาวิทยาลัยแม่โจ้, รายงานผลการวิจัย, มี.ค. 2556.

ภาคผนวก ก

ข้อมูลคุณลักษณะของกุ้ง

ตารางที่ ก-1 ข้อมูลคุณลักษณะของตัวอย่างกุ้งจำนวน 200 ภาพ

ตัวอย่าง	Area	Length	Mean	Std	Min	Per25	Per50	Per75	Group
1	6.40	79.18	158.97	33.41	22.00	150.25	303.50	784.50	1
2	6.32	68.07	157.69	32.00	22.00	128.25	265.00	900.00	1
3	5.01	59.46	158.87	29.20	27.00	113.25	248.50	1082.50	1
4	3.84	47.27	160.99	27.20	26.00	111.00	241.00	795.25	1
5	2.96	41.24	161.49	25.20	29.00	99.00	223.00	400.00	1
6	1.93	29.36	163.41	21.88	30.00	73.25	180.00	962.25	1
7	1.39	26.76	165.53	19.20	37.00	64.00	175.00	600.00	1
8	1.25	18.08	165.01	20.10	30.00	54.25	190.00	715.00	1
9	0.60	5.79	166.53	18.45	43.00	18.50	133.00	807.50	2
10	5.75	60.00	158.00	29.00	20.00	40.00	143.50	1051.75	1
11	4.79	69.64	159.61	28.47	31.00	131.50	302.00	786.25	1
12	4.57	68.40	158.45	26.49	31.00	151.25	292.00	934.25	1
13	4.38	63.82	158.06	25.20	24.00	123.75	260.00	898.25	1
14	2.60	49.90	160.66	23.80	30.00	106.00	242.50	940.50	1
15	1.89	34.87	161.08	21.10	32.00	54.00	201.00	888.75	1
16	1.50	21.25	159.42	19.10	31.00	38.25	143.50	873.25	1
17	5.50	13.26	158.00	28.00	21.00	41.25	117.00	907.75	1
18	4.75	50.00	157.00	14.00	28.00	50.00	115.00	1235.50	1
19	3.00	53.46	163.00	21.00	34.00	70.00	170.00	769.25	1
20	0.67	22.24	160.00	18.00	42.00	10.25	145.50	1051.75	2
21	0.77	19.39	167.93	15.60	47.00	6.75	104.50	587.00	2
22	0.50	9.82	167.00	16.60	45.00	2.25	136.00	500.00	2
23	1.75	35.00	163.00	22.00	30.00	130.00	132.50	991.00	1
24	4.00	48.46	158.78	24.00	24.00	150.00	190.50	1106.00	1
25	3.00	33.20	161.45	23.00	25.00	90.00	169.00	991.50	1
26	2.75	18.88	163.23	22.00	32.00	69.00	210.00	981.50	1
27	2.50	24.33	162.94	21.00	36.00	90.00	148.00	1083.75	1
28	0.80	13.53	167.00	19.60	49.00	12.25	191.00	1513.75	2
29	0.60	12.00	164.00	18.60	42.00	4.25	143.00	1422.50	2
30	2.75	54.07	166.30	21.26	32.00	50.50	220.00	1568.00	1
31	4.04	49.93	165.95	24.00	23.00	75.50	179.00	1212.00	1
32	3.25	33.95	167.09	23.00	28.00	56.50	133.50	1321.00	1

ตารางที่ ก-1 ข้อมูลคุณลักษณะของตัวอย่างกึ่งจำนวน 200 ภาพ(ต่อ)

ตัวอย่าง	Area	Length	Mean	Std	Min	Per25	Per50	Per75	Group
33	3.50	19.40	168.71	23.00	32.00	24.75	210.00	1430.00	1
34	0.75	5.15	169.60	17.20	45.00	13.00	55.00	722.00	2
35	0.25	0.63	170.52	16.70	51.00	1.00	97.50	1310.25	2
36	3.71	70.97	159.33	22.80	31.00	27.00	226.50	1109.75	1
37	4.08	63.60	159.40	20.50	30.00	27.00	259.50	1166.00	1
38	2.09	49.90	158.72	20.50	36.00	19.25	265.00	1212.50	1
39	2.59	36.73	156.82	20.10	33.00	16.25	224.00	1146.00	1
40	3.35	31.00	158.28	19.40	29.00	89.00	197.00	1307.00	1
41	1.89	17.46	160.60	19.00	38.00	0.00	175.50	1081.00	1
42	0.48	10.37	161.98	18.04	48.00	6.00	45.00	1110.50	2
43	5.50	65.00	161.79	27.00	38.00	77.00	142.50	1656.00	1
44	5.75	69.40	154.89	30.74	31.00	69.75	359.50	799.00	1
45	4.69	61.42	154.70	27.50	25.00	90.50	313.50	934.00	1
46	3.87	49.34	158.44	25.14	26.00	80.75	258.00	706.00	1
47	2.80	31.57	161.45	22.80	29.00	58.00	216.00	670.75	1
48	1.44	15.61	168.09	22.00	30.00	23.00	168.00	775.75	1
49	0.55	6.13	163.84	19.10	49.00	17.25	80.00	999.00	2
50	3.75	46.00	151.00	26.00	26.00	15.25	259.00	1368.00	1
51	2.77	62.46	159.03	25.44	28.00	105.25	248.50	1720.50	1
52	2.14	46.67	160.93	24.70	24.00	115.00	237.50	1122.25	1
53	2.53	40.20	161.07	23.20	30.00	75.50	180.50	1203.50	1
54	1.25	20.66	164.93	20.50	31.00	44.50	420.00	1019.25	1
55	0.75	6.60	166.42	17.00	39.00	17.00	81.50	712.00	2
56	2.75	38.00	153.00	22.00	35.00	4.00	450.00	597.25	1
57	5.00	72.92	158.00	30.21	24.00	147.50	385.50	948.25	1
58	4.95	73.10	158.90	27.20	29.00	182.00	404.50	1298.25	1
59	4.00	63.91	159.93	25.90	25.00	168.25	365.00	1374.75	1
60	3.00	57.78	161.06	24.50	29.00	127.25	323.00	1133.00	1
61	3.31	49.23	162.31	23.30	28.00	85.50	278.00	1132.00	1
62	2.79	38.10	161.78	21.90	29.00	91.75	252.00	1077.50	1
63	1.75	25.47	163.09	21.00	26.00	110.00	229.50	1118.75	1
64	1.50	14.43	164.43	20.00	35.00	52.00	169.50	1045.25	1

ตารางที่ ก-1 ข้อมูลคุณลักษณะของตัวอย่างกุ้งจำนวน 200 ภาพ(ต่อ)

ตัวอย่าง	Area	Length	Mean	Std	Min	Per25	Per50	Per75	Group
65	0.70	10.00	165.35	17.20	44.00	26.50	103.00	1068.25	2
66	0.60	8.00	168.41	15.25	45.00	12.00	82.00	1001.25	2
67	5.94	56.87	154.20	27.80	27.00	171.25	455.00	1700.00	1
68	6.09	59.13	152.96	27.30	11.00	185.75	414.50	1600.00	1
69	4.53	49.18	154.59	25.40	12.00	154.25	376.50	1322.75	1
70	3.76	37.26	157.56	23.40	23.00	129.75	343.00	1367.25	1
71	2.66	25.46	156.93	22.34	32.00	46.75	350.50	1500.00	1
72	2.10	16.05	158.85	20.76	32.00	43.25	260.00	1800.00	1
73	0.76	6.47	161.07	19.40	45.00	18.00	193.50	1900.00	2
74	0.10	0.00	164.00	17.45	39.00	14.25	85.00	1137.25	2
75	3.18	65.20	155.39	23.13	27.00	80.25	256.50	1400.00	1
76	2.84	61.32	155.78	24.00	26.00	111.00	240.00	1696.25	1
77	3.22	50.34	157.50	22.60	27.00	105.00	238.50	1600.00	1
78	2.32	37.22	155.26	21.10	35.00	35.75	229.50	1934.00	1
79	1.36	25.50	157.45	19.50	28.00	19.25	170.50	1795.25	1
80	0.73	8.69	160.00	17.70	42.00	8.25	105.00	1337.00	2
81	0.60	8.00	162.00	18.00	46.00	2.25	95.00	1800.00	2
82	1.59	33.89	163.86	23.00	29.00	83.00	203.50	1196.75	1
83	1.58	32.75	162.19	20.00	25.00	67.75	168.00	1120.75	1
84	1.74	22.85	166.39	19.00	28.00	57.00	153.00	719.00	1
85	0.75	13.65	164.24	17.50	40.00	4.25	116.00	1047.75	2
86	0.50	12.00	163.70	17.00	41.00	2.00	102.00	960.00	2
87	0.30	9.00	166.17	16.00	45.00	3.00	95.00	773.75	2
88	6.00	70.02	144.45	30.00	14.00	130.00	420.00	962.75	1
89	6.11	59.77	154.30	30.00	20.00	114.75	386.50	864.00	1
90	4.30	44.11	154.88	26.00	21.00	112.75	338.50	905.00	1
91	4.87	43.59	157.00	27.00	21.00	116.50	325.50	833.25	1
92	4.02	31.22	161.00	24.00	24.00	93.25	216.00	650.00	1
93	1.20	11.90	164.80	20.00	32.00	51.50	219.00	551.50	1
94	1.20	11.00	166.00	15.00	48.00	4.00	86.00	606.00	2
95	6.33	68.29	155.00	31.00	14.00	86.00	357.00	805.50	1
96	4.15	59.20	157.00	28.00	20.00	73.25	324.00	895.50	1

ตารางที่ ก-1 ข้อมูลคุณลักษณะของตัวอย่างกึ่งจำนวน 200 ภาพ(ต่อ)

ตัวอย่าง	Area	Length	Mean	Std	Min	Per25	Per50	Per75	Group
97	3.66	46.34	159.90	26.00	24.00	86.50	246.50	561.75	1
98	3.94	28.28	158.31	23.00	26.00	73.00	236.50	300.00	1
99	1.99	22.26	160.45	21.00	29.00	89.25	184.00	830.50	1
100	5.75	55.00	150.00	33.00	16.00	37.00	200.00	829.25	1
101	2.50	42.00	159.00	23.00	30.00	37.00	250.00	740.00	1
102	5.50	55.00	148.00	29.00	18.00	84.00	420.00	777.00	1
103	5.00	49.00	150.00	28.00	26.00	83.00	300.00	1118.75	1
104	4.50	45.00	152.00	27.00	27.00	109.00	220.00	1045.25	1
105	3.00	29.00	153.00	23.00	29.00	44.00	290.00	1068.25	2
106	2.00	23.00	158.00	21.00	34.00	31.00	240.00	1001.25	2
107	1.75	19.00	164.00	21.00	34.00	25.00	190.00	1700.00	1
108	0.50	12.00	169.00	15.50	42.00	5.00	58.00	1600.00	1
109	5.25	60.00	150.00	29.00	21.00	90.00	399.00	1322.75	1
110	5.25	50.00	151.00	27.00	21.00	88.00	400.00	1367.25	1
111	4.50	35.00	154.00	26.50	25.00	78.00	324.00	1500.00	1
112	2.75	27.00	159.00	23.00	34.00	43.00	299.00	1800.00	1
113	2.25	25.00	158.00	22.00	34.00	21.00	300.00	1900.00	2
114	1.00	14.00	171.00	19.00	42.00	3.00	148.00	1137.25	2
115	3.25	33.00	158.00	23.00	28.00	105.00	190.00	1400.00	1
116	5.25	47.00	152.00	24.00	19.00	90.00	234.00	1696.25	1
117	4.50	41.00	154.00	24.00	18.00	65.00	312.00	690.00	1
118	3.25	33.00	156.00	22.00	31.00	39.00	322.00	547.00	1
119	3.25	30.00	155.00	23.00	26.00	38.00	270.00	1082.50	1
120	2.50	22.00	164.00	20.00	33.00	27.00	289.00	1350.00	1
121	2.25	21.00	163.00	22.00	34.00	22.00	311.00	1200.00	1
122	3.75	38.00	157.00	28.00	28.00	36.00	301.00	962.25	1
123	6.25	69.00	146.00	35.00	15.00	95.00	379.00	1400.00	1
124	5.50	54.00	147.00	31.00	17.00	79.00	366.00	715.00	1
125	4.75	48.00	152.00	29.00	38.00	71.00	354.00	600.00	1
126	3.50	38.00	159.00	26.00	32.00	57.00	283.00	1051.75	1
127	3.50	40.00	159.00	26.00	33.00	54.00	294.00	1022.00	1
128	2.50	34.00	163.00	24.00	33.00	45.00	255.00	1067.00	1



ตารางที่ ก-1 ข้อมูลคุณลักษณะของตัวอย่างกุ้งจำนวน 200 ภาพ(ต่อ)

ตัวอย่าง	Area	Length	Mean	Std	Min	Per25	Per50	Per75	Group
129	0.75	14.00	166.00	22.00	37.00	8.00	90.00	579.00	2
130	5.25	49.00	151.00	30.00	18.00	119.00	347.00	1082.50	1
131	3.50	31.00	157.00	27.00	26.00	54.00	254.00	1149.00	1
132	3.25	28.00	158.00	25.00	28.00	49.00	243.00	720.00	1
133	2.25	63.00	160.00	19.00	34.00	56.00	381.00	962.25	1
134	0.91	24.00	168.00	18.00	37.00	15.00	120.00	480.00	2
135	2.75	27.00	162.00	25.00	31.00	47.00	389.00	715.00	1
136	5.50	68.00	154.00	32.00	21.00	82.00	345.00	1190.00	1
137	4.00	43.00	154.00	29.00	26.00	80.00	296.00	1051.75	1
138	3.50	49.00	156.00	28.00	36.00	74.00	312.00	436.00	1
139	2.50	38.00	160.00	26.00	25.00	67.00	326.00	1120.00	1
140	2.25	62.00	162.00	23.00	26.00	65.00	285.00	1256.00	1
141	0.20	14.00	168.00	18.00	42.00	20.00	79.25	690.25	2
142	4.50	38.00	154.00	29.00	20.00	90.00	399.00	1200.00	1
143	3.75	38.00	155.00	26.00	22.00	88.00	400.00	690.00	1
144	3.50	34.00	157.00	27.00	26.00	78.00	324.00	745.00	1
145	2.50	28.00	163.00	21.00	33.00	43.00	299.00	1090.00	1
146	2.75	33.00	162.00	23.00	30.00	54.25	157.50	715.00	1
147	0.75	6.00	169.00	17.30	41.00	7.00	145.50	1058.75	2
148	3.25	41.00	156.00	26.00	30.00	67.00	190.00	440.00	1
149	5.25	56.00	145.00	31.00	18.00	90.00	234.00	1345.00	1
150	3.50	38.00	152.00	29.00	23.00	65.00	312.00	690.00	1
151	3.50	39.00	150.00	27.00	26.00	39.00	322.00	547.00	1
152	2.75	33.00	156.00	26.00	24.00	38.00	270.00	1082.50	1
153	2.25	26.00	157.00	22.00	30.00	27.00	289.00	1350.00	1
154	1.75	19.00	162.00	21.00	36.00	47.00	389.00	715.00	1
155	5.00	61.00	148.00	32.00	17.00	82.00	345.00	1190.00	1
156	3.50	42.00	150.00	28.00	25.00	80.00	296.00	1051.75	1
157	3.75	44.00	151.00	29.00	26.00	74.00	312.00	436.00	1
158	3.25	38.00	156.00	27.00	26.00	67.00	326.00	1120.00	1
159	2.50	31.00	156.00	23.00	39.00	90.00	234.00	1345.00	1
160	0.80	8.00	171.00	18.00	40.00	2.00	100.00	532.00	2

ตารางที่ ก-1 ข้อมูลคุณลักษณะของตัวอย่างกุ้งจำนวน 200 ภาพ(ต่อ)

ตัวอย่าง	Area	Length	Mean	Std	Min	Per25	Per50	Per75	Group
161	5.50	67.00	148.00	29.00	16.00	40.00	143.50	1051.75	1
162	4.50	54.00	152.00	27.00	21.00	131.50	302.00	786.25	1
163	4.25	48.00	152.00	26.00	26.00	151.25	292.00	934.25	1
164	2.75	28.00	161.00	22.00	39.00	123.75	275.50	898.25	1
165	2.25	27.00	161.00	22.00	38.00	106.00	242.50	940.50	1
166	1.75	29.00	163.00	19.00	37.00	54.00	201.00	888.75	1
167	4.75	48.00	153.00	27.00	23.00	38.25	143.50	873.25	1
168	3.00	40.00	160.00	24.00	33.00	41.25	117.00	907.75	1
169	2.50	28.00	161.00	22.00	37.00	65.00	115.00	1235.50	1
170	2.00	22.00	162.00	20.00	37.00	70.00	135.50	769.25	1
171	5.75	69.00	150.00	34.00	16.00	90.00	399.00	1200.00	1
172	4.50	52.00	157.00	29.00	20.00	88.00	400.00	690.00	1
173	3.50	38.00	164.00	27.00	26.00	78.00	324.00	745.00	1
174	3.25	34.00	164.00	25.00	24.00	43.00	299.00	1090.00	1
175	3.25	36.00	164.00	24.00	27.00	50.00	115.00	1235.50	1
176	1.75	24.00	162.00	20.00	38.00	70.00	135.50	769.25	1
177	0.00	0.00	172.00	15.00	43.00	11.00	120.00	1481.00	2
178	5.50	67.00	146.00	32.00	16.00	83.00	203.50	1196.75	1
179	4.25	54.00	151.00	31.00	26.00	67.75	168.00	1120.75	1
180	4.25	44.00	150.00	30.00	26.00	57.00	153.00	719.00	1
181	0.20	23.00	164.00	19.00	36.00	9.00	67.00	1002.25	2
182	4.75	41.00	153.00	29.00	21.00	40.00	143.50	1051.75	1
183	4.75	39.00	152.00	28.00	21.00	131.50	302.00	786.25	1
184	3.25	36.00	157.00	28.00	22.00	151.25	292.00	934.25	1
185	3.25	37.00	156.00	27.00	29.00	123.75	275.50	898.25	1
186	2.00	25.00	163.00	26.00	32.00	106.00	242.50	940.50	1
187	1.25	18.00	164.00	20.50	36.00	110.00	201.00	888.75	1
188	2.50	28.00	162.00	23.00	33.00	38.25	143.50	873.25	1
189	6.25	72.00	142.00	33.00	14.00	41.25	117.00	907.75	1
190	5.00	59.00	150.00	31.00	16.00	50.00	50.00	1235.50	1
191	4.00	56.00	156.00	27.00	22.00	70.00	135.50	769.25	1
192	2.75	38.00	161.00	24.00	31.00	83.00	203.50	1196.75	1

ตารางที่ ก-1 ข้อมูลคุณลักษณะของตัวอย่างกุ้งจำนวน 200 ภาพ(ต่อ)

ตัวอย่าง	Area	Length	Mean	Std	Min	Per25	Per50	Per75	Group
193	2.75	29.00	160.00	24.00	31.00	67.75	168.00	1120.75	1
194	1.50	27.00	164.00	21.00	32.00	57.00	153.00	719.00	1
195	1.00	21.00	166.00	19.00	37.00	24.00	75.00	1104.75	2
196	5.50	62.00	153.00	28.00	14.00	90.00	399.00	1200.00	1
197	4.00	53.00	160.00	25.00	22.00	88.00	400.00	690.00	1
198	3.75	39.00	163.00	24.00	30.00	78.00	324.00	745.00	1
199	3.75	42.00	164.00	24.00	26.00	43.00	299.00	1090.00	1
200	2.75	22.00	165.00	21.00	33.00	54.25	157.50	715.00	1

หมายเหตุ : Area แทนค่า สัดส่วนพื้นที่  
 Length แทนค่า สัดส่วนความยาว  
 Mean แทนค่า ค่าเฉลี่ย  
 Std แทนค่า ส่วนเบี่ยงเบนมาตรฐาน  
 Min แทนค่า ค่าความสว่างต่ำสุด  
 Per 25 แทนค่า จำนวนพิกเซล ณ เปอร์เซ็นไทล์ที่ 25  
 Per 50 แทนค่า จำนวนพิกเซล ณ เปอร์เซ็นไทล์ที่ 50  
 Per 75 แทนค่า จำนวนพิกเซล ณ เปอร์เซ็นไทล์ที่ 75  
 Group แทนค่า กลุ่มมาตรฐาน 1 แทน กลุ่มไม่ผ่านมาตรฐาน 2 แทน กลุ่มผ่านมาตรฐาน

ภาคผนวก ข

การเตรียมข้อมูลก่อนการจำแนก

### การทดสอบค่าเฉลี่ยกรณีกลุ่มตัวอย่าง 2 กลุ่ม

พิจารณาจากการนำคุณลักษณะมาทดสอบทางสถิติ t-test จากกลุ่มตัวอย่าง ประชากร 2 กลุ่ม โดยนำค่าเฉลี่ยของแต่ละคุณลักษณะมาทำการทดสอบความแตกต่างของกลุ่มทั้ง ผ่านและไม่ผ่านเกณฑ์มาตรฐานซึ่งก่อนนำมาทดสอบด้วย t-test ต้องทำการทดสอบเงื่อนไขความ เท่ากันของความแปรปรวนด้วย F-test โดยตั้งสมมติฐานดังตารางที่ 3.1 ที่ระดับความเชื่อมั่น 95% ( $\alpha=0.05$ ) ดังนี้

ทดสอบ F-test

$H_0$  : ความแปรปรวนของคุณลักษณะแต่ละประเภทไม่แตกต่างกัน

$H_1$  : ความแปรปรวนของคุณลักษณะแต่ละประเภทแตกต่างกัน

ทดสอบ T-test

$H_0$  : ค่าเฉลี่ยของคุณลักษณะแต่ละประเภทไม่แตกต่างกัน

$H_1$  : ค่าเฉลี่ยของคุณลักษณะแต่ละประเภทแตกต่างกัน

หลังจากนั้นทำการทดสอบความแปรปรวนโดยใช้สถิติ F-test ได้ผลลัพธ์การ วิเคราะห์ด้วยโปรแกรม SPSS เวอร์ชัน 22 ดังตาราง ข-1

ตาราง ข-1 Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means		
		F	Sig.	t	df	Sig. (2-tailed)
Area	Equal variances assumed	46.172	.000	12.003	198	.000
	Equal variances not assumed			25.472	196.455	.000
Length	Equal variances assumed	29.676	.000	10.884	198	.000
	Equal variances not assumed			18.908	112.556	.000
Mean	Equal variances assumed	7.835	.006	-8.617	198	.000
	Equal variances not assumed			-11.824	62.696	.000

ตาราง ข-1 Independent Samples Test (ต่อ)

		Levene's Test for Equality of Variances		t-test for Equality of Means		
		F	Sig.	t	df	Sig. (2-tailed)
Std	Equal variances assumed	27.704	.000	10.610	198	.000
	Equal variances not assumed			18.326	110.544	.000
Min	Equal variances assumed	7.708	.006	-13.644	198	.000
	Equal variances not assumed			-18.650	62.254	.000
Per25	Equal variances assumed	33.607	.000	9.831	198	.000
	Equal variances not assumed			21.129	197.847	.000
Per50	Equal variances assumed	25.151	.000	9.844	198	.000
	Equal variances not assumed			16.760	105.744	.000
Per75	Equal variances assumed	2.382	.124	.189	198	.850
	Equal variances not assumed			.162	37.398	.872

ผลลัพธ์จากการประมวลผลตาราง ข-1 เมื่อทำการทดสอบเงื่อนไขความเท่ากันของความแปรปรวนด้วยการทดสอบสถิติ F-test ที่มีค่านัยสำคัญเท่ากับ 0.05 ปรากฏว่าค่า Sig. ของ F-test ของ 7 คุณลักษณะ คือ พื้นที่ ความยาว ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าต่ำสุด เปอร์เซนต์ไทล์ที่ 25 และ 50 น้อยกว่า 0.05 จึงปฏิเสธสมมติฐาน  $H_0$  แสดงว่าความแปรปรวนของทั้งสองกลุ่มนี้แตกต่างกันอย่างมีนัยสำคัญ จึงใช้ค่าสถิติ t-test ที่ Equal Variances not assumed ซึ่งมีการพิจารณาลักษณะเช่นเดียวกับกับ F-test โดย t-test ปรากฏว่าค่า Sig. ของทุกคุณลักษณะข้างต้นมีค่าน้อยกว่า 0.05 เช่นกัน จึงปฏิเสธสมมติฐาน  $H_0$  แสดงว่าค่าเฉลี่ยของคุณลักษณะมีความแตกต่างกันอย่างมีนัยสำคัญ และเปอร์เซนต์ไทล์ที่ 75 ทดสอบเงื่อนไขความเท่ากันของความแปรปรวนด้วยการทดสอบสถิติ F-test ที่มีค่านัยสำคัญเท่ากับ 0.05 ปรากฏว่าค่า Sig. ของ F-test มากกว่า 0.05 จึงยอมรับสมมติฐาน  $H_0$  แสดงว่าความแปรปรวนของทั้งสองกลุ่มนี้ไม่แตกต่างกันอย่างมีนัยสำคัญ จึงใช้ค่าสถิติ t-test ที่ Equal Variances assumed ซึ่งมีการพิจารณาลักษณะเช่นเดียวกับกับ F-test โดย t-test ปรากฏว่าค่า Sig. ของทุกคุณลักษณะข้างต้นมีค่ามากกว่า 0.05 จึงยอมรับสมมติฐาน  $H_0$  แสดงว่าค่าเฉลี่ยของคุณลักษณะไม่มีความแตกต่างกันอย่างมีนัยสำคัญ ดังนั้นคุณลักษณะที่จะนำมาใช้เป็น

พารามิเตอร์ในการจำแนกกลุ่มมี 7 คุณลักษณะ คือ พื้นที่ ความยาว ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าต่ำสุด จำนวนพิกเซล ณ เปอร์เซนต์ไทล์ที่ 25 และ 50

### การวิเคราะห์ปัจจัยเพื่อลดขนาดข้อมูล

ในงานวิจัยนำคุณลักษณะที่ส่งต่อผล มาลดขนาดข้อมูลด้วยเทคนิค Principal Component Analysis (PCA) ของตัวอย่างภาพทั้งจำนวน 200 ภาพ จะได้ผลลัพธ์การวิเคราะห์ด้วยโปรแกรม SPSS เวอร์ชัน 22 ดังต่อไปนี้

จากตาราง ข-2 เป็นค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson correlation) จะพบว่าพารามิเตอร์ที่มีความสัมพันธ์กันมากที่สุด คือ พื้นที่ และส่วนเบี่ยงเบนมาตรฐาน โดยมีค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.881 ดังนั้นพารามิเตอร์ของพื้นที่ และส่วนเบี่ยงเบนมาตรฐานควรจัดอยู่ในปัจจัยเดียวกัน

ตาราง ข-2 เมทริกซ์สหสัมพันธ์ (Correlation matrix)

	Area	Length	Mean	Std	Min	Per25	Per50
Area	1.000	.843	-.766	.881	-.832	.613	.588
Length	.843	1.000	-.659	.773	-.729	.631	.554
Mean	-.766	-.659	1.000	-.769	.721	-.404	-.522
Std	.881	.773	-.769	1.000	-.802	.576	.579
Min	-.832	-.729	.721	-.802	1.000	-.595	-.558
Per25	.613	.631	-.404	.576	-.595	1.000	.546
Per50	.588	.554	-.522	.579	-.558	.546	1.000

จากตาราง ข-3 ค่า Kaiser-Meyer-Olkin Measure of Sampling เป็นการวัดความเหมาะสมของข้อมูล ในการใช้เทคนิคลดมิติข้อมูลนี้ โดยมีค่าเท่ากับ 0.915 ซึ่งมากกว่า 0.50 และใกล้กับ 1 ดังนั้นข้อมูลที่มีอยู่เหมาะสมที่จะใช้เทคนิคนี้ และใช้สถิติทดสอบ Bartlett's Test of Sphericity โดยกำหนดสมมติฐาน ดังนี้

$H_0$ : พารามิเตอร์ต่างๆ ไม่มีความสัมพันธ์กัน

$H_1$ : พารามิเตอร์ต่าง ๆ มีความสัมพันธ์กัน

สถิติทดสอบจะมีการแจกแจงโดยประมาณแบบ Chi-Square เท่ากับ 1214.174 และค่า Sig. เท่ากับ 0.000 ซึ่งมีค่าน้อยกว่า 0.05 จึงปฏิเสธ  $H_0$  นั่นคือพารามิเตอร์มีความสัมพันธ์กัน ดังนั้นสามารถใช้เทคนิคนี้ในการวิเคราะห์ต่อไปได้

ตาราง ข-3 สถิติ KMO และการทดสอบของ Bartlett

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.915
Bartlett's Test of Sphericity	Approx. Chi-Square	1214.174
	df	21
	Sig.	.000

จากตาราง ข-4 แสดงค่าสถิติสำหรับแต่ละปัจจัยทั้งก่อนและหลังการสกัดปัจจัย โดยวิธี PCA ในการสกัดปัจจัย ซึ่งมีรายละเอียดดังนี้

ปัจจัย (Component) เป็นปัจจัยที่จัดกลุ่มซึ่งจะมีจำนวนปัจจัยมากที่สุดเท่ากับจำนวนพารามิเตอร์ โดยในขั้นตอนนี้มี 7 พารามิเตอร์ จึงสามารถจัดกลุ่มได้มากที่สุด 7 ปัจจัยหรือ 7 Component

ค่าไอเกน (Eigenvalues) เป็นค่าความผันแปรหรือความแปรปรวนทั้งหมดในพารามิเตอร์เดิมที่สามารถอธิบายได้ด้วยปัจจัย โดยการจัดพารามิเตอร์จะไม่พิจารณาปัจจัยที่มีค่าไอเกนน้อยกว่า 1 ซึ่งเป็นค่าที่นิยมใช้จากตารางจะพบว่ามีเพียงปัจจัยที่ 1 เท่านั้นที่มีค่าไอเกนมากกว่า 1 ดังนั้นการจัดพารามิเตอร์จึงมีแค่ 1 ปัจจัยเท่านั้น และจากกราฟ scree plot ในรูปที่ ข-1 เป็นกราฟที่พล็อตค่าไอเกนของแต่ละปัจจัยเรียงค่าจากมากไปน้อย เพื่อใช้ในการพิจารณาประกอบการพิจารณาค่าจากการคำนวณ ซึ่งจะพิจารณาจากค่าไอเกนที่ลดลงอย่างรวดเร็ว และมีค่าไอเกนมากกว่า 1 จากรูปจึงพิจารณาว่ามี 1 ปัจจัย แต่เมื่อพิจารณาข้อมูลจากการคำนวณพบว่ามีบางปัจจัยที่มีค่าไอเกนใกล้เคียงกับ 1 และจากการศึกษาพบว่านอกจากค่าไอเกนดังกล่าว ยังมีค่าไอเกนที่ใช้ในการวิเคราะห์การจัดพารามิเตอร์ นั่นคือ 0.7 หรือมากกว่าทำให้ผู้วิจัยตัดสินใจเพิ่มปัจจัย เพื่อเพิ่มเปอร์เซ็นต์ที่จะสามารถอธิบายความแปรปรวนได้ นอกจากนั้นทำให้สามารถวิเคราะห์ด้วยการหมุนแกนปัจจัย ในกรณีที่มีการแปลผล และค่า Factor loading มีค่าไม่ชัดเจนในการจัดพารามิเตอร์ว่าควรจัดอยู่ในปัจจัยใด จึงเพิ่มปัจจัยที่ 2 ในการวิเคราะห์ โดยมีค่าไอเกนเท่ากับ 0.676 ซึ่งใกล้เคียงกับค่าที่กำหนด

% of Variance เป็นเปอร์เซ็นต์ที่แต่ละปัจจัยสามารถอธิบายความแปรปรวนได้ เนื่องจากเดิมมีพารามิเตอร์ 7 ตัว และจากตารางจะพบว่าแต่ละตัวมีค่า Communalities เท่ากับ 1 จึงมีความแปรปรวนทั้งหมดเท่ากับ 7



ดังนั้น % of Variance ของปัจจัยที่ 1 =  $(5.027/7)*100 = 71.819\%$

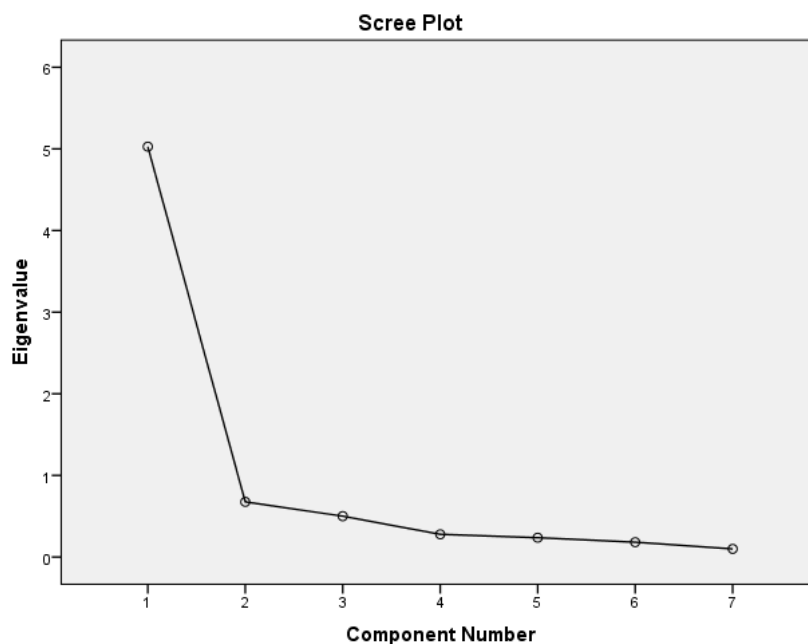
% of Variance ของปัจจัยที่ 2 =  $(0.676/7)*100 = 9.651\%$

โดย Cumulative % หมายถึง ผลบวกสะสมของความแปรปรวน จึงสามารถสรุปได้ว่าปัจจัยที่สำคัญที่สุดคือ ปัจจัยที่ 1 เนื่องจากสามารถอธิบายความผันแปรของข้อมูลได้มากที่สุดคือเท่ากับ 71.819 ส่วน ปัจจัยที่ 2 จะมีส่วนสำคัญรองลงมาสามารถอธิบายความแปรปรวนของข้อมูลได้ 9.651 ส่วนรวมสองปัจจัยเป็น 81.471 ส่วน

ในส่วนของการ Rotation Sums of Squared Loadings จะให้ค่าไอเกน เพอร์เซ็นต์ ความแปรปรวน และ ผลบวกสะสมของความแปรปรวนของปัจจัยเมื่อทำการหมุนแกนปัจจัยในลักษณะที่ปัจจัยยังคงเป็นอิสระต่อกัน ในการวิเคราะห์นี้ใช้วิธีหมุนแกนปัจจัยแบบ Varimax จะพบว่าค่าต่างๆของปัจจัยที่ 1 เมื่อหมุนแกนมีค่าน้อยกว่าเมื่อไม่ได้หมุนแกน ในขณะที่ปัจจัยที่ 2 มีค่าไอเกน และเปอร์เซ็นต์ความแปรปรวน มากกว่าปัจจัยที่ 1 แต่ผลบวกสะสมของความแปรปรวน ของทั้ง 2 ปัจจัยยังคงเท่าเดิม

ตาราง ข-4 ความแปรปรวนทั้งหมดที่สามารถอธิบายได้ (Total Variance Explained)

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.027	71.819	71.819	5.027	71.819	71.819	3.547	50.673	50.673
2	.676	9.651	81.471	.676	9.651	81.471	2.156	30.798	81.471
3	.501	7.153	88.623						
4	.278	3.975	92.599						
5	.237	3.381	95.980						
6	.182	2.595	98.575						
7	.100	1.425	100.000						



รูป ข-1 กราฟ scree plot

จากตาราง ข-5 แสดงค่าสัมประสิทธิ์สหสัมพันธ์ (Factor loading) ซึ่งเป็นความสัมพันธ์ของพารามิเตอร์กับปัจจัยทั้ง 2 ปัจจัยโดยไม่หมุนแกนปัจจัย นั่นคือการทำให้อัตราตั้งฉากกันหรือเป็นอิสระกัน โดยให้ทั้ง 2 ปัจจัยตั้งฉากกันหรือเป็นอิสระกันได้ผลลัพธ์ดังตาราง ข-6 ตาราง ข-5 เมทริกซ์ค่าถ่วงน้ำหนักปัจจัยก่อนหมุนแกนปัจจัย (Component Matrix<sup>a</sup>)

	Component	
	1	2
Area	.942	
Std	.918	
Min	-.892	
Length	.882	
Mean	-.825	.384
Per25	.725	.564
Per50	.720	.390

เมื่อหมุนแกนปัจจัยโดยวิธี Varimax พบว่าค่าสัมประสิทธิ์สหสัมพันธ์มีค่าเปลี่ยนแปลงไปเมื่อเทียบกับไม่หมุนแกน ทำให้บางปัจจัยมีค่าสัมประสิทธิ์สหสัมพันธ์มากขึ้น ดังนั้นสามารถจัดปัจจัยได้ดังนี้

ปัจจัยที่ 1 ประกอบด้วยตัวแปร 5 พารามิเตอร์คือ ค่าเฉลี่ย (Mean), ส่วนเบี่ยงเบนมาตรฐาน (Std), พื้นที่ (Area), ค่าความสว่างที่น้อยที่สุด (Min) และความยาว (Length)

ปัจจัยที่ 2 ประกอบด้วยตัวแปร 2 พารามิเตอร์คือ ค่าเปอร์เซ็นต์ไทล์ที่ 25 (Per25) และค่าเปอร์เซ็นต์ไทล์ที่ 50 (Per50)

ตาราง ข-6 เมทริกซ์ค่าถ่วงน้ำหนักปัจจัยหลังจากหมุนแกนปัจจัย (Rotated Component Matrix<sup>a</sup>)

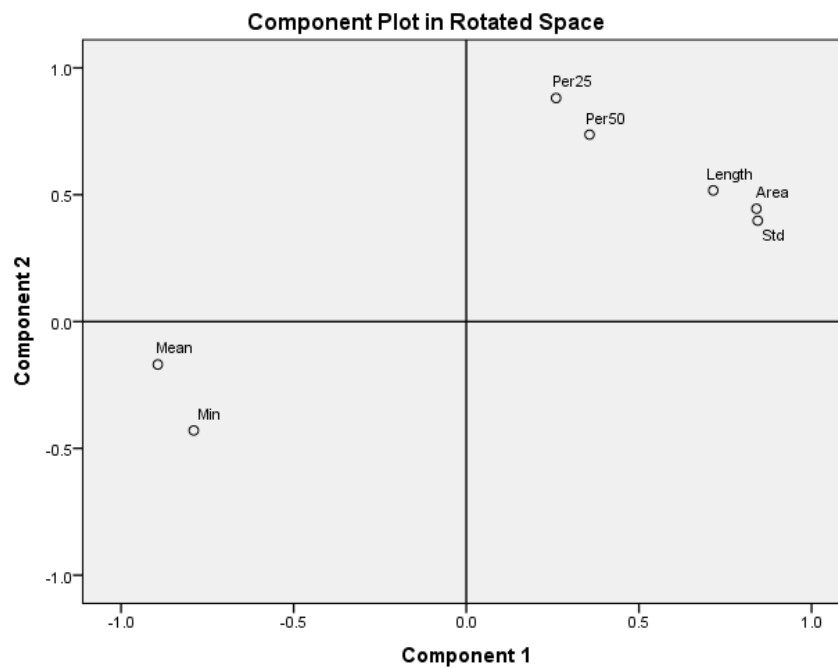
	Component	
	1	2
Mean	-.894	
Std	.844	.398
Area	.840	.445
Min	-.790	-.429
Length	.715	.517
Per25	.260	.881
Per50	.357	.737

ค่าของเมทริกซ์การหมุนที่ใช้ในการหมุนแกนปัจจัยเพื่อเปลี่ยนค่าถ่วงน้ำหนัก ภายหลังจากการหมุนแกนปัจจัยแล้วโดยใช้วิธี Varimax ดังตาราง ข-7

ตาราง ข-7 เมทริกซ์การหมุนแกนปัจจัยเพื่อเปลี่ยนค่าถ่วงน้ำหนักปัจจัย

Component	1	2
1	.812	.583
2	-.583	.812

จากรูป ข-2 แสดงค่าถ่วงน้ำหนักปัจจัยของแต่ละปัจจัย โดยที่ถ้าปัจจัยสามารถแทน พารามิเตอร์ต่างๆได้ดี พารามิเตอร์จะต้องอยู่ใกล้กับแกนใดแกนหนึ่ง ถ้าพารามิเตอร์อยู่ใกล้จุด กำเนิด (0,0,0) แสดงว่าพารามิเตอร์นั้นไม่สัมพันธ์กับปัจจัยใดเลย ในที่นี้พารามิเตอร์ทั้ง 7 พารามิเตอร์อยู่บริเวณปลายแกน จึงจัดพารามิเตอร์ให้อยู่ในปัจจัยได้ และพารามิเตอร์ที่อยู่ในปัจจัย เดียวกันมีความสัมพันธ์กันมาก



รูป ข-2 Component plot in rotated space

จากพารามิเตอร์ 7 พารามิเตอร์สามารถจัดให้เหลือ 2 ปัจจัยเหมือนเป็นการลดพารามิเตอร์จาก 7 พารามิเตอร์เหลือ 2 พารามิเตอร์ โดยถือว่าปัจจัยเป็นพารามิเตอร์ใหม่ ซึ่งค่าคะแนนปัจจัยหาได้จากสมการ 3.4 โดยจากตาราง ข-6 สามารถคำนวณหาค่าคะแนนปัจจัยได้ดังนี้  
ค่าคะแนนปัจจัยที่ 1 คือ

$$F_{11} = -0.894Z_{11} + 0.844Z_{21} + 0.84Z_{31} - 0.79Z_{41} + 0.715Z_{51} + 0.26Z_{61} + 0.357Z_{71} \quad (\text{ข-1})$$

ค่าคะแนนปัจจัยที่ 2 คือ

$$F_{21} = 0.398Z_{22} + 0.445Z_{32} - 0.429Z_{42} + 0.517Z_{52} + 0.881Z_{62} + 0.737Z_{72} \quad (\text{ข-2})$$

ภาคผนวก ค

การจำแนกกลุ่มด้วยเทคนิค Discriminant Analysis

## การจำแนกกลุ่มด้วยเทคนิค Discriminant Analysis

### 1. กรณีข้อมูลไม่มีการลดขนาด

โดยขั้นตอนนี้ต้องตรวจสอบข้อสมมติฐานเบื้องต้นก่อนนำข้อมูลไปวิเคราะห์ประกอบด้วย 2 ประการดังนี้

- พารามิเตอร์หรือคุณลักษณะทั้ง 7 ตัวต้องมีการแจกแจงแบบปกติตัวแปรหลายตัว ในการตรวจสอบการแจกแจงแบบปกติตัวแปรหลายตัวนั้น ให้ตรวจสอบการแจกแจงแบบปกติของตัวแปรครั้งละตัว ถ้าพบว่าพารามิเตอร์แต่ละตัวมีการแจกแจงแบบปกติ ไม่สามารถสรุปได้ว่าพารามิเตอร์ทั้ง 7 ตัวจะมีการแจกแจงแบบปกติตัวแปรหลายตัว แต่ถ้าพารามิเตอร์ตัวใดตัวหนึ่งไม่มีการแจกแจงแบบปกติ สรุปได้ว่าพารามิเตอร์ทั้ง 7 ตัว ไม่มีการแจกแจงแบบปกติตัวแปรหลายตัว และพารามิเตอร์ทั้ง 7 ตัว มีการแจกแจงแบบปกติตัวแปรหลายตัว สรุปได้ว่าพารามิเตอร์แต่ละตัวมีการแจกแจงแบบปกติ ดังนั้นจึงตรวจสอบพารามิเตอร์แต่ละตัวถึงแม้ว่าเมื่อพารามิเตอร์ทุกตัวมีการแจกแจงแบบปกติแล้ว จะไม่สามารถสรุปได้ว่าพารามิเตอร์ทั้ง 7 ตัว มีการแจกแจงแบบปกติตัวแปรหลายตัวก็ตาม แต่มีความน่าจะเป็นสูงที่จะมีการแจกแจงแบบปกติหลายตัวแปร และถ้าตรวจสอบแล้วพบว่าพารามิเตอร์บางตัวไม่มีการแจกแจงแบบปกติ แสดงว่าพารามิเตอร์ทั้ง 7 ตัว ไม่ได้มีการแจกแจงแบบปกติตัวแปรหลายตัว

จากตาราง ค-1 พบว่าพารามิเตอร์ที่มีค่า Sig มากกว่า 0.05 คือ พื้นที่ ความยาว ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และจำนวนพิทเชลสะสม ณ เปอร์เซ็นไทล์ที่ 50 แสดงว่าพารามิเตอร์ทั้ง 5 ตัวมีการแจกแจงแบบปกติ ส่วนพารามิเตอร์ที่มีค่า Sig มีค่าน้อยกว่าค่าระดับนัยสำคัญ นั่นคือ น้อยกว่า 0.05 คือ ค่าต่ำสุด และจำนวนพิทเชลสะสม ณ เปอร์เซ็นไทล์ที่ 25 ไม่มีการแจกแจงแบบปกติ ดังนั้นจึงไม่นำพารามิเตอร์ดังกล่าวมาใช้ในการจำแนกกลุ่มด้วยวิธี discriminant analysis

ตาราง ค-1 One-Sample Kolmogorov-Smirnov Test

		Area	Length	Mean	Std	Min	Per25	Per50
N		140	140	140	140	140	140	140
Normal Parameters <sup>a,b</sup>	Mean	3.0212	37.3045	159.6816	23.2833	30.3357	64.8107	238.4679
	Std. Deviation	1.6991	19.8210	5.2490	4.3817	8.3223	44.8049	96.6559
Most Extreme Differences	Absolute	.062	.065	.060	.069	.092	.079	.047
	Positive	.062	.060	.028	.069	.092	.079	.047
	Negative	-.051	-.065	-.060	-.045	-.055	-.074	-.032
Test Statistic		.062	.065	.060	.069	.092	.079	.047
Asymp. Sig. (2-tailed)		.200 <sup>c,d</sup>	.200 <sup>c,d</sup>	.200 <sup>c,d</sup>	.200 <sup>c,d</sup>	.005 <sup>c</sup>	.031 <sup>c</sup>	.200 <sup>c,d</sup>

- เมทริกซ์ความแปรปรวนร่วมของพารามิเตอร์ในแต่ละกลุ่มต้องเท่ากัน โดยใช้สถิติทดสอบ Box's M

จากตาราง ค-2 ค่า Sig ของการทดสอบเท่ากับ 0.18 ซึ่งมีค่ามากกว่าค่าระดับนัยสำคัญ นั่นคือ มีค่ามากกว่า 0.05 ดังนั้นเมทริกซ์ความแปรปรวนร่วมของพารามิเตอร์ของกลุ่มที่ 1 และกลุ่มที่ 2 ไม่แตกต่างกัน สรุปได้ว่าข้อมูลเป็นไปตามข้อสมมติเบื้องต้นทั้ง 2 ข้อของการจำแนกประเภท จึงสามารถนำข้อมูลไปทำการวิเคราะห์จำแนกต่อไป

ตาราง ค-2 Test Results

Box's M		107.032
F	Approx.	6.573
	df1	15
	df2	7328.299
	Sig.	.180

จากตาราง ค-3 แสดงความแปรปรวน และความแปรปรวนร่วมของพารามิเตอร์แยกตามกลุ่ม ถ้าความแปรปรวนร่วมของพารามิเตอร์คู่ใดมีค่ามาก แสดงว่าพารามิเตอร์คู่นั้นมีความสัมพันธ์กันมาก โดยในกลุ่มที่ 1 และ กลุ่มที่ 2 พารามิเตอร์ความยาว และ เปอร์เซ็นต์ไทล์ที่ 50 เท่ากับ 579.452 และ 57.853 ซึ่งมีค่ามากแสดงว่าพารามิเตอร์ทั้งสองมีความสัมพันธ์กันมาก

ตาราง ค-3 Covariance Matrices

Group		Area	Length	Mean	Std	Per50
1	Area	2.001	16.475	-4.391	4.363	51.468
	Length	16.475	280.371	-35.339	39.439	579.452
	Mean	-4.391	-35.339	22.263	-10.476	-213.467
	Std	4.363	39.439	-10.476	14.253	153.577
	Per50	51.468	579.452	-213.467	153.577	6527.716
2	Area	.054	.694	.070	.045	2.324
	Length	.694	32.395	-.676	.815	57.853
	Mean	.070	-.676	9.548	-1.292	-17.104
	Std	.045	.815	-1.292	2.473	23.975
	Per50	2.324	57.853	-17.104	23.975	1389.652

จากตาราง ค-4 แสดงค่าลักษณะเฉพาะ ซึ่งในการจำแนกกลุ่มข้อมูลมี 2 กลุ่ม จึงมีเพียงสมการเดียว และมีค่าลักษณะเฉพาะค่าเดียว ซึ่งมีค่าเท่ากับ 1.038 และมีค่าสหสัมพันธ์คาโนนิคอล เท่ากับ 0.714 เมื่อนำค่าดังกล่าวมายกกำลังสอง จะเป็นค่าที่แสดงให้ทราบว่าตัวแปรในสมการจำแนกกลุ่ม สามารถอธิบายความผันแปรของตัวแปรตามได้ร้อยละ  $(0.714)^2 = 0.5098$  หรือเท่ากับ 50.98%

ตาราง ค-4 Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	1.038 <sup>a</sup>	100.0	100.0	.714

จากตาราง ค-5 Wilks' Lambda เป็นค่าที่แสดงสัดส่วนความแปรผันของคะแนนการจำแนกกลุ่มที่ไม่สามารถอธิบายได้ด้วยความแตกต่างระหว่างกลุ่ม สถิติทดสอบวิธีนี้ใช้ทดสอบการเท่ากันของค่าเฉลี่ยพารามิเตอร์ทุกตัวในกลุ่มที่ 1 และกลุ่มที่ 2 โดยค่า Wilks' Lambda เท่ากับ 0.491 และค่า Sig เท่ากับ 0.000 ซึ่งมีค่าน้อยกว่าค่าระดับนัยสำคัญ ดังนั้นค่าเฉลี่ยของพารามิเตอร์ทุกตัวในกลุ่มที่ 1 ไม่เท่ากับกลุ่มที่ 2 ซึ่งแสดงว่าสมการที่ได้จากการวิเคราะห์จำแนกกลุ่มสามารถจำแนกออกเป็น 2 กลุ่มได้

ตาราง ค-5 Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.491	96.448	5	.000

จากตาราง ค-6 แสดงค่าสัมประสิทธิ์ (ค่าถ่วงน้ำหนัก) ของพารามิเตอร์จำแนกในสมการจำแนกกลุ่มซึ่งเป็นสมการมาตรฐาน (Standardized) เป็นค่าถ่วงน้ำหนักของพารามิเตอร์แต่ละตัว จึงไม่มีหน่วย โดยพารามิเตอร์พื้นที่มีค่าถ่วงน้ำหนักเท่ากับ 0.502 ซึ่งมีค่ามากกว่าค่าถ่วงน้ำหนักของพารามิเตอร์ส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ -0.047 แสดงว่าพารามิเตอร์พื้นที่มีความสำคัญในการจำแนกกลุ่มในสมการจำแนกกลุ่มมากกว่าพารามิเตอร์ส่วนเบี่ยงเบนมาตรฐาน ซึ่งสามารถเขียนเป็นสมการจำแนกกลุ่มได้ดังนี้

$$Z_Y = 0.502Z_{Area} + 0.299Z_{Length} + 0.086Z_{Mean} - 0.047Z_{std} + 0.524Z_{Per50} \quad (ค-1)$$



ตาราง ค-6 Standardized Canonical Discriminant Function Coefficients

	Function
	1
Area	.502
Length	.299
Mean	.086
Std	-.047
Per50	.524

จากตาราง ค-7 เป็นการประเมินพารามิเตอร์แต่ละตัวส่งผลต่อการจำแนกอย่างไร เป็นการพิจารณาความสัมพันธ์ระหว่างค่าของพารามิเตอร์แต่ละตัวกับคะแนนจำแนกกลุ่ม จะพบว่าพารามิเตอร์พื้นที่ที่มีความสำคัญต่อการจำแนกกลุ่มมากที่สุด รองลงมาคือพารามิเตอร์จำนวนพิกเซล ณ เปอร์เซนต์ที่ 50 ความยาว และส่วนเบี่ยงเบนมาตรฐานตามลำดับ ส่วนพารามิเตอร์ค่าเฉลี่ยมีความสำคัญต่อการจำแนกกลุ่มน้อย

ตาราง ค-7 Structure Matrix

	Function 1
Area	.850
Per50	.804
Length	.802
Std	.751
Mean	-.608

จากตาราง ค-8 แสดงค่าเป็นสัมประสิทธิ์ (ค่าถ่วงน้ำหนัก) ของพารามิเตอร์จำแนกในสมการจำแนกกลุ่มที่อยู่ในรูปคะแนนดิบ จึงไม่อยู่ในรูปมาตรฐาน (Unstandardized) ผลที่ได้มีทั้งค่าน้ำหนักในแต่ละตัวแปร และค่าคงที่ (Constant) จากผลสามารถเขียนเป็นสมการจำแนกกลุ่มในรูปคะแนนดิบได้ดังนี้

ตาราง ค-8 Canonical Discriminant Function Coefficients

	Function
	1
Area	.390
Length	.019
Mean	.019
Std	-.013
Per50	.007
(Constant)	-6.319

Unstandardized coefficients

$$\hat{D} = 6.319 + .390\text{Area} + .019\text{Length} + .019\text{Mean} - 0.013\text{Std} + 0.007\text{Per50} \quad (\text{ค-2})$$

จากตาราง ค-9 แสดงค่ากลางของกลุ่ม (Group Centroids) โดยกลุ่มที่ 1 มีค่ากลางของกลุ่มเท่ากับ 0.472 ส่วนกลุ่มที่ 2 มีค่ากลางของกลุ่มเท่ากับ -2.169 ซึ่งค่ากลางของทั้งสองกลุ่มแตกต่างกัน แสดงว่าสมการดังกล่าวสามารถจำแนกกลุ่มได้ดี ในกรณีที่มีค่าวิเคราะห์ใหม่สามารถคำนวณคะแนนจำแนกกลุ่ม แล้วนำมาเปรียบเทียบกับค่ากลางของแต่ละกลุ่ม ถ้าค่าเข้าใกล้ค่ากลางกลุ่มใดจะจัดให้ค่านั้นเป็นสมาชิกกลุ่มนั้น

ตาราง ค-9 Functions at Group Centroids

Group	Function
	1
1	.472
2	-2.169

Unstandardized canonical discriminant functions evaluated at group means

ตาราง ค-10 เป็นผลการวิเคราะห์ค่าสัมประสิทธิ์ (ค่าถ่วงน้ำหนัก) และค่าคงที่ของสมการจำแนกกลุ่มโดยแยกเป็นกลุ่มตามวิธี Fisher's linear discrimination function จำนวนสมการจะมีเท่ากับจำนวนกลุ่ม ในกรณีนี้มี 2 กลุ่มจึงมี 2 สมการ จากผลการวิเคราะห์จะได้สมการดังนี้ สมการของกลุ่ม 1 :

$$\hat{D} = -1429.122 + 21.577\text{Area} - 0.426\text{Length} + 16.325\text{Mean} + 5.229\text{Std} + 0.318\text{Per50} \quad (\text{ค-2})$$

สมการของกลุ่ม 2

$$\hat{D} = -1414.677 + 20.548\text{Area} - 0.477\text{Length} + 16.274\text{Mean} + 5.264\text{Std} + 0.300\text{Per50} \quad (\text{ค-3})$$

ตาราง ค-10 Classification Function Coefficients

	Group	
	1	2
Area	21.577	20.548
Length	-.426	-.477
Mean	16.325	16.274
Std	5.229	5.264
Per50	.318	.300
(Constant)	-1429.122	-1414.677

Fisher's linear discriminant functions

จากผลการวิเคราะห์สามารถนำไปใช้พยากรณ์ข้อมูลใหม่ เพื่อให้ทราบว่าควรจัดให้อยู่กลุ่มใด โดยการแทนค่าพารามิเตอร์ จากสมการคือ พื้นที่ ความยาว ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และจำนวนพิทเชลสะสม ณ เปอร์เซนต์ที่ 50 ของข้อมูลภาพนั้นๆ ลงในทั้ง 2 สมการ ถ้าสมการของกลุ่มที่ 1 มีค่ามากกว่าสมการของกลุ่มที่ 2 จะจัดให้ภาพดังกล่าวอยู่ในกลุ่มที่ 1 แต่ถ้าสมการของกลุ่มที่ 2 มีค่ามากกว่าสมการของกลุ่มที่ 1 จะจัดให้ภาพดังกล่าวอยู่ในกลุ่มที่ 2 นั่นคือ ถ้าสมการใดมีค่ามากกว่าก็จัดอยู่ในกลุ่มนั้น

## 2. กรณีข้อมูลมีการลดขนาด

โดยขั้นตอนนี้ต้องตรวจสอบข้อสมมติฐานเบื้องต้นก่อนนำข้อมูลไปวิเคราะห์ ประกอบด้วย 2 ประการดังนี้

- พารามิเตอร์หรือคุณลักษณะทั้ง 2 ตัวต้องมีการแจกแจงแบบปกติตัวแปรหลายตัว การตรวจสอบพารามิเตอร์แต่ละตัวแม้ว่าเมื่อพารามิเตอร์ทุกตัวมีการแจกแจงแบบปกติแล้ว จะไม่สามารถสรุปได้ว่าพารามิเตอร์ทั้ง 2 ตัว มีการแจกแจงแบบปกติตัวแปรหลายตัวก็ตาม แต่มีความน่าจะเป็นสูงที่จะมีการแจกแจงแบบปกติหลายตัวแปร และถ้าตรวจสอบแล้วพบว่าพารามิเตอร์บางตัวไม่มีการแจกแจงแบบปกติ แสดงว่าพารามิเตอร์ทั้ง 2 ตัว ไม่ได้มีการแจกแจงแบบปกติตัวแปรหลายตัว จากตาราง ค-11 พบว่าพารามิเตอร์ทั้ง 2 ตัวมีค่า Sig มากกว่า 0.05

ตาราง ค-11 One-Sample Kolmogorov-Smirnov Test

		fac1_1	fac2_1
N		140	140
Normal Parameters <sup>a,b</sup>	Mean	-.1102860	.0494645
	Std. Deviation	.87215181	1.01689515
Most Extreme Differences	Absolute	.066	.046
	Positive	.066	.046
	Negative	-.044	-.039
Test Statistic		.066	.046
Asymp. Sig. (2-tailed)		.200 <sup>c,d</sup>	.200 <sup>c,d</sup>

- เมทริกซ์ความแปรปรวนร่วมของพารามิเตอร์ในแต่ละกลุ่มต้องเท่ากัน โดยใช้สถิติทดสอบ Box's M

จากตาราง ค-12 ค่า Sig ของการทดสอบเท่ากับ 0.25 ซึ่งมีความมากกว่าค่าระดับนัยสำคัญ นั่นคือ มีค่ามากกว่า 0.05 ดังนั้นเมทริกซ์ความแปรปรวนร่วมของพารามิเตอร์ของกลุ่มที่ 1 และกลุ่มที่ 2 ไม่แตกต่างกัน

ตาราง ค-12 Test Results

Box's M		60.138
F	Approx.	19.418
	df1	3
	df2	24924.082
	Sig.	.250

สรุปได้ว่าข้อมูลเป็นไปตามข้อสมมติเบื้องต้นทั้ง 2 ข้อ ของการจำแนกประเภท จึงสามารถนำข้อมูลไปทำการวิเคราะห์จำแนกต่อไป

จากตาราง ค-13 แสดงความแปรปรวน และความแปรปรวนร่วมของพารามิเตอร์แยกตามกลุ่ม ถ้าความแปรปรวนร่วมของพารามิเตอร์คู่ใดมีค่ามาก แสดงว่าพารามิเตอร์คู่นั้นมีความสัมพันธ์กันมาก

ตาราง ค-13 Covariance Matrices

Group		fac1_1	fac2_1
1	fac1_1	.614	-.095
	fac2_1	-.095	.800
2	fac1_1	.106	-.055
	fac2_1	-.055	.085

จากตาราง ค-14 แสดงค่าลักษณะเฉพาะซึ่งในการจำแนกกลุ่มข้อมูลมี 2 กลุ่ม จึงมีเพียงสมการเดียว และมีค่าลักษณะเฉพาะค่าเดียว ซึ่งมีค่าเท่ากับ 1.173 และมีค่าสหสัมพันธ์คาโนนิคอล เท่ากับ 0.735 เมื่อนำค่าดังกล่าวมายกกำลังสอง จะเป็นค่าที่แสดงให้ทราบว่าตัวแปรในสมการจำแนกกลุ่ม สามารถอธิบายความผันแปรของตัวแปรตามได้ร้อยละเท่าไร ซึ่งกรณีนี้อธิบายได้ร้อยละ  $(0.735)^2 = 0.5402$  หรือเท่ากับ 54.02%

ตาราง ค-14 Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	1.173 <sup>a</sup>	100.0	100.0	.735

จากตาราง ค-15 Wilks' Lambda เป็นค่าที่แสดงสัดส่วนความแปรผันของคะแนนการจำแนกกลุ่มที่ไม่สามารถอธิบายได้ด้วยความแตกต่างระหว่างกลุ่ม สถิติทดสอบวิธีนี้ใช้ทดสอบการเท่ากันของค่าเฉลี่ยพารามิเตอร์ทุกตัวในกลุ่มที่ 1 และกลุ่มที่ 2 โดยค่า Wilks' Lambda เท่ากับ .460 และค่า Sig เท่ากับ 0.000 ซึ่งมีค่าน้อยกว่าค่าระดับนัยสำคัญ ดังนั้นค่าเฉลี่ยของพารามิเตอร์ทุกตัวในกลุ่มที่ 1 ไม่เท่ากับกลุ่มที่ 2 ซึ่งแสดงว่าสมการที่ได้จากการวิเคราะห์จำแนกกลุ่มสามารถจำแนกออกเป็น 2 กลุ่มได้

ตาราง ค-15 Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.460	106.304	2	.000

จากตาราง ค-16 แสดงค่าสัมประสิทธิ์ (ค่าถ่วงน้ำหนัก) ของพารามิเตอร์จำแนกในสมการจำแนกกลุ่มซึ่งเป็นสมการมาตรฐาน (Standardized) เป็นค่าถ่วงน้ำหนักของพารามิเตอร์แต่ละตัว จึงไม่มีหน่วย โดยพารามิเตอร์พื้นที่มีค่าถ่วงน้ำหนักเท่ากับ 0.502 ซึ่งมีค่ามากกว่าค่าถ่วงน้ำหนักของพารามิเตอร์ส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ -0.047 แสดงว่าพารามิเตอร์พื้นที่มีความสำคัญในการจำแนกกลุ่มในสมการจำแนกกลุ่มมากกว่าพารามิเตอร์ส่วนเบี่ยงเบนมาตรฐาน ซึ่งสามารถเขียนเป็นสมการจำแนกกลุ่มได้ดังนี้

$$Z_Y = 0.742Z_{fac1} + 0.789Z_{fac2} \quad (\text{ค-4})$$

ตาราง ค-16 Standardized Canonical Discriminant Function Coefficients

	Function
	1
fac1_1	.742
fac2_1	.789

จากตาราง ค-17 เป็นการประเมินพารามิเตอร์แต่ละตัวส่งผลต่อการจำแนกอย่างไร เป็นการพิจารณาความสัมพันธ์ระหว่างค่าของพารามิเตอร์แต่ละตัวกับคะแนนจำแนกกลุ่ม จะพบว่าพารามิเตอร์ปัจจัยที่ 1 มีความสำคัญต่อการจำแนกกลุ่มมากที่สุด และรองลงมาคือพารามิเตอร์ปัจจัยที่ 2

ตาราง ค-17 Structure Matrix

	Function 1
fac1_1	.679
fac2_1	.625

จากตาราง ค-18 แสดงค่าเป็นสัมประสิทธิ์ (ค่าถ่วงน้ำหนัก) ของพารามิเตอร์จำแนกในสมการจำแนกกลุ่มที่อยู่ในรูปคะแนนดิบ จึงไม่อยู่ในรูปมาตรฐาน (Unstandardized) ผลที่ได้มีทั้งค่าน้ำหนักในแต่ละตัวแปร และค่าคงที่ (Constant) จากผลสามารถเขียนเป็นสมการจำแนกกลุ่มในรูปคะแนนดิบได้ดังนี้

$$\hat{D} = 0.065 + 1.024fac1\_1 + 0.960fac2\_1 \quad (\text{ค-5})$$

ตาราง ค-18 Canonical Discriminant Function Coefficients

	Function
	1
fac1_1	1.024
fac2_1	.960
(Constant)	.065

จากตาราง ค-19 แสดงค่ากลางของกลุ่ม (Group Centroids) โดยกลุ่มที่ 1 มีค่ากลางของกลุ่มเท่ากับ 0.501 ส่วนกลุ่มที่ 2 มีค่ากลางของกลุ่มเท่ากับ -2.306 ซึ่งค่ากลางของทั้งสอง

กลุ่มแตกต่างกัน แสดงว่าสมการดังกล่าวสามารถจำแนกกลุ่มได้ดี ในกรณีที่มีค่าวิเคราะห์ใหม่สามารถคำนวณคะแนนจำแนกกลุ่ม แล้วนำมาเปรียบเทียบกับค่ากลางของแต่ละกลุ่ม ถ้าค่าเข้าใกล้ค่ากลางกลุ่มใดจะจัดให้ค่านั้นเป็นสมาชิกกลุ่มนั้น

ตาราง ค-19 Functions at Group Centroids

Group	Function
	1
1	.501
2	-2.306

Unstandardized canonical discriminant functions evaluated at group means

ตาราง ค-20 เป็นผลการวิเคราะห์ค่าสัมประสิทธิ์ (ค่าถ่วงน้ำหนัก) และค่าคงที่ของสมการจำแนกกลุ่มโดยแยกเป็นกลุ่มตามวิธี Fisher's linear discrimination function จำนวนสมการจะมีเท่ากับจำนวนกลุ่ม ในกรณีนี้มี 2 กลุ่มจึงมี 2 สมการ จากผลการวิเคราะห์จะได้สมการดังนี้ สมการของกลุ่ม 1 :

$$\hat{D} = -0.798 + 0.528\text{fac1\_1} - 0.798\text{fac2\_1} \quad (\text{ค-6})$$

สมการของกลุ่ม 2

$$\hat{D} = -3.515 - 2.562\text{fac1\_1} - 2.167\text{fac2\_1} \quad (\text{ค-7})$$

ตาราง ค-20 Classification Function Coefficients

	Group	
	1	2
fac1_1	.311	-2.562
fac2_1	.528	-2.167
(Constant)	-.798	-3.515

Fisher's linear discriminant functions

จากผลการวิเคราะห์สามารถนำไปใช้พยากรณ์ข้อมูลใหม่ เพื่อให้ทราบว่าควรจัดให้อยู่กลุ่มใด โดยการแทนค่าพารามิเตอร์ลงในทั้ง 2 สมการ ถ้าสมการของกลุ่มที่ 1 มีค่ามากกว่าสมการของกลุ่มที่ 2 จะจัดให้ภาพดังกล่าวอยู่ในกลุ่มที่ 1 แต่ถ้าสมการของกลุ่มที่ 2 มีค่ามากกว่าสมการของกลุ่มที่ 1 จะจัดให้ภาพดังกล่าวอยู่ในกลุ่มที่ 2 นั่นคือ ถ้าสมการใดมีค่ามากกว่าก็จัดอยู่ในกลุ่มนั้น

ภาคผนวก ง

การจำแนกกลุ่มด้วยเทคนิคซ์พอร์ทเวคเตอร์แมชชีน



## การจำแนกกลุ่มด้วยเทคนิคซ์พอร์ดเวกเตอร์แมชชีน

### 1. กรณีข้อมูลไม่มีการลดขนาด

- ฟังก์ชันคอร์เนลแบบเชิงเส้น

การทดสอบประสิทธิภาพความแม่นยำของอัลกอริทึมซ์พอร์ดเวกเตอร์แมชชีนแบ่งข้อมูลในการทดสอบประสิทธิภาพด้วยวิธี self-consistency test , split test และ leave-one-out cross-validation ผลการนำข้อมูลสำหรับเรียนรู้และสร้างโมเดลด้วยวิธีซ์พอร์ดเวกเตอร์แมชชีนในกรณีของข้อมูลภาพไม่ผ่านกระบวนการลดขนาดด้วยใช้โปรแกรม WEKA ผลจากการทดสอบโมเดลวิธีดังกล่าวแสดงดังรูป ง-1, ง-2 และ ง-3 ตามลำดับ

```

=== Summary ===
Correctly Classified Instances      200          100.0 %
Incorrectly Classified Instances    0            0.0 %
Kappa statistic                     1
Mean absolute error                 0
Root mean squared error             0
Relative absolute error             0 %
Root relative squared error         0 %
Total Number of Instances          200

=== Confusion Matrix ===
  a  b  <-- classified as
169  0 | a = 1
  0 31 | b = 2

```

รูป ง-1 ผลการวัดประสิทธิภาพของโมเดลด้วยวิธีซ์พอร์ดเวกเตอร์แมชชีนฟังก์ชันคอร์เนลแบบเชิงเส้นแบ่งข้อมูลด้วยวิธี self-consistency test

```

=== Summary ===
Correctly Classified Instances      56      93.3  %
Incorrectly Classified Instances    4       6.7  %
Kappa statistic                    0.6154
Mean absolute error                0.1
Root mean squared error            0.3162
Relative absolute error            35.1485 %
Root relative squared error        78.2064 %
Total Number of Instances          60

=== Confusion Matrix ===
 a  b  <-- classified as
48  0 | a = 1
 4  8 | b = 2

```

รูป ง-2 ผลการวัดประสิทธิภาพของโมเดลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนฟังก์ชันเคอร์เนล  
แบบเชิงเส้นแบ่งข้อมูลด้วยวิธี split test

```

=== Summary ===
Correctly Classified Instances      192     96.0  %
Incorrectly Classified Instances    8       4.0  %
Kappa statistic                    0.9812
Mean absolute error                0.005
Root mean squared error            0.0707
Relative absolute error            1.8872 %
Root relative squared error        19.533 %
Total Number of Instances          200

=== Confusion Matrix ===
 a  b  <-- classified as
169  0 | a = 1
 8  23 | b = 2

```

รูป ง-3 ผลการวัดประสิทธิภาพของโมเดลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนฟังก์ชันเคอร์เนล  
แบบเชิงเส้นแบ่งข้อมูลด้วยวิธี leave one out cross-validation

- ฟังก์ชันเคอร์เนลแบบโพลีโนเมียล

การทดสอบประสิทธิภาพความแม่นยำของอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีนแบ่งข้อมูลในการทดสอบประสิทธิภาพด้วยวิธี self-consistency test , split test และ leave-one-out cross-validation ผลการนำข้อมูลสำหรับเรียนรู้และสร้างโมเดลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนในกรณีของข้อมูลภาพไม่ผ่านกระบวนการลดขนาดด้วยโปรแกรม WEKA ผลจากการทดสอบโมเดลวิธีดังกล่าวแสดงดังรูป ง-4, ง-5 และ ง-6 ตามลำดับ

```

=== Summary ===
Correctly Classified Instances      200          100.0 %
Incorrectly Classified Instances     0            0.0 %
Kappa statistic                      1
Mean absolute error                  0
Root mean squared error              0
Relative absolute error              0 %
Root relative squared error          0 %
Total Number of Instances           200

=== Confusion Matrix ===
  a  b  <-- classified as
169  0 | a = 1
  0  31 | b = 2

```

รูป ง-4 ผลการวัดประสิทธิภาพของโมเดลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนฟังก์ชันเคอร์เนลแบบโพลีโนเมียลแบ่งข้อมูลด้วยวิธี self-consistency test

```

=== Summary ===
Correctly Classified Instances      48      80.0  %
Incorrectly Classified Instances    12      20.0  %
Kappa statistic                    0
Mean absolute error                0.2
Root mean squared error            0.4472
Relative absolute error            70.297  %
Root relative squared error        110.6005 %
Total Number of Instances          60

=== Confusion Matrix ===
 a  b  <-- classified as
48  0 | a = 1
12  0 | b = 2

```

รูป ง-5 ผลการวัดประสิทธิภาพของโมเดลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนฟังก์ชันเคอร์เนล  
แบบโพลีโนเมียลแบ่งข้อมูลด้วยวิธี split test

```

=== Summary ===
Correctly Classified Instances      169     84.5  %
Incorrectly Classified Instances     31     15.5  %
Kappa statistic                    0
Mean absolute error                0.155
Root mean squared error            0.3937
Relative absolute error            58.3536 %
Root relative squared error        108.2423 %
Total Number of Instances          200

=== Confusion Matrix ===
 a  b  <-- classified as
169  0 | a = 1
31  0 | b = 2

```

รูป ง-6 ผลการวัดประสิทธิภาพของโมเดลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนฟังก์ชันเคอร์เนล  
แบบโพลีโนเมียลแบ่งข้อมูลด้วยวิธี leave one out cross-validation

## 2. กรณีข้อมูลไม่มีการลดขนาด

- ฟังก์ชันเคอร์เนลแบบเชิงเส้น

การทดสอบประสิทธิภาพความแม่นยำของอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีนแบ่งข้อมูลในการทดสอบประสิทธิภาพด้วยวิธี self-consistency test , split test และ leave-one-out cross-validation ผลการนำข้อมูลสำหรับเรียนรู้และสร้างโมเดลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนในกรณีของข้อมูลภาพผ่านกระบวนการลดขนาดด้วยโปรแกรม WEKA ผลจากการทดสอบโมเดลวิธีดังกล่าวแสดงดังรูป ง-7, ง-8 และ ง-9 ตามลำดับ

```

=== Summary ===
Correctly Classified Instances      200      100.0 %
Incorrectly Classified Instances     0         0.0 %
Kappa statistic                     1
Mean absolute error                 0
Root mean squared error             0
Relative absolute error             0 %
Root relative squared error         0 %
Total Number of Instances          200

=== Confusion Matrix ===
 a  b  <-- classified as
169  0 | a = 1
 0 31 | b = 2

```

รูป ง-7 ผลการวัดประสิทธิภาพของโมเดลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนฟังก์ชันเคอร์เนลแบบเชิงเส้นแบ่งข้อมูลด้วยวิธี self-consistency test

```

=== Summary ===
Correctly Classified Instances      48      80.0  %
Incorrectly Classified Instances    12      20.0  %
Kappa statistic                     0
Mean absolute error                 0.2
Root mean squared error             0.4472
Relative absolute error             70.297  %
Root relative squared error         110.6005 %
Total Number of Instances          60

=== Confusion Matrix ===
 a  b  <-- classified as
48  0 | a = 1
12  0 | b = 2

```

รูป ง-8 ผลการวัดประสิทธิภาพของโมเดลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนฟังก์ชันเคอร์เนล  
แบบเชิงเส้นแบ่งข้อมูลด้วยวิธี split test

```

=== Summary ===
Correctly Classified Instances      169     84.5  %
Incorrectly Classified Instances     31     15.5  %
Kappa statistic                     0
Mean absolute error                 0.155
Root mean squared error             0.3937
Relative absolute error             58.3536 %
Root relative squared error         108.2423 %
Total Number of Instances          200

=== Confusion Matrix ===
 a  b  <-- classified as
169  0 | a = 1
31  0 | b = 2

```

รูป ง-9 ผลการวัดประสิทธิภาพของโมเดลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนฟังก์ชันเคอร์เนล  
แบบเชิงเส้นแบ่งข้อมูลด้วยวิธี leave one out cross-validation

- ฟังก์ชันเคอร์เนลแบบโพลีโนเมียล

การทดสอบประสิทธิภาพความแม่นยำของอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีนแบ่งข้อมูลในการทดสอบประสิทธิภาพด้วยวิธี self-consistency test , split test และ leave-one-out cross-validation ผลการนำข้อมูลสำหรับเรียนรู้และสร้างโมเดลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนในกรณีของข้อมูลภาพผ่านกระบวนการลดขนาดด้วยโปรแกรม WEKA ผลจากการทดสอบโมเดลวิธีดังกล่าวแสดงดังรูป ง-10, ง-11 และ ง-12 ตามลำดับ

```

=== Summary ===
Correctly Classified Instances      200          100.0 %
Incorrectly Classified Instances    0             0.0 %
Kappa statistic                     1
Mean absolute error                 0
Root mean squared error             0
Relative absolute error             0 %
Root relative squared error         0 %
Total Number of Instances          200

=== Confusion Matrix ===
 a  b  <-- classified as
169  0 | a = 1
 0  31 | b = 2

```

รูป ง-10 ผลการวัดประสิทธิภาพของโมเดลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนฟังก์ชันเคอร์เนลแบบโพลีโนเมียลแบ่งข้อมูลด้วยวิธี self-consistency test

```

=== Summary ===
Correctly Classified Instances      48      80.0  %
Incorrectly Classified Instances    12      20.0  %
Kappa statistic                     0
Mean absolute error                 0.2
Root mean squared error             0.4472
Relative absolute error             70.297  %
Root relative squared error         110.6005 %
Total Number of Instances          60

=== Confusion Matrix ===
 a  b  <-- classified as
48  0 | a = 1
12  0 | b = 2

```

รูป ง-11 ผลการวัดประสิทธิภาพของโมเดลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนฟังก์ชันเคอร์เนล  
แบบโพลีโนเมียลแบ่งข้อมูลด้วยวิธี split test

```

=== Summary ===
Correctly Classified Instances      169     84.5  %
Incorrectly Classified Instances    31     15.5  %
Kappa statistic                     0
Mean absolute error                 0.155
Root mean squared error             0.3937
Relative absolute error             58.5042 %
Root relative squared error         108.7551 %
Total Number of Instances          200

=== Confusion Matrix ===
 a  b  <-- classified as
169  0 | a = 1
31  0 | b = 2

```

รูป ง-12 ผลการวัดประสิทธิภาพของโมเดลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนฟังก์ชันเคอร์เนล  
แบบโพลีโนเมียลแบ่งข้อมูลด้วยวิธี leave one out cross-validation



