

การรู้จำชนิดของเสื้อผ้า โดยใช้วิธีการจำแนกประเภท



บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)  
เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ ที่ส่งผ่านทางบัณฑิตวิทยาลัย

The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository (CUIR)  
are the thesis authors' files submitted through the University Graduate School.

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต  
สาขาวิชาวิศวกรรมซอฟต์แวร์ ภาควิชาวิศวกรรมคอมพิวเตอร์  
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย  
ปีการศึกษา 2557  
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

Recognition Types of Clothing using Classification Method

Mr. Wisarut Surakarin



A Thesis Submitted in Partial Fulfillment of the Requirements  
for the Degree of Master of Science Program in Software Engineering

Department of Computer Engineering

Faculty of Engineering

Chulalongkorn University

Academic Year 2014

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์ การรู้จำชนิดของเส้นผ้า โดยใช้วิธีการจำแนกประเภท  
โดย นายวิศรุต สุรการินทร์  
สาขาวิชา วิศวกรรมซอฟต์แวร์  
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก ศาสตราจารย์ ดร. ประภาส จงสฤษดิ์วัฒนา

---

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้บัณฑิตวิทยานิพนธ์ฉบับนี้เป็นส่วน  
หนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรบัณฑิต

.....คณบดีคณะวิศวกรรมศาสตร์  
(ศาสตราจารย์ ดร. บัณฑิต เอื้ออาภรณ์)

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ  
(รองศาสตราจารย์ ดร. วิวัฒน์ วัฒนาวุฒิ)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก  
(ศาสตราจารย์ ดร. ประภาส จงสฤษดิ์วัฒนา)

.....กรรมการภายนอกมหาวิทยาลัย  
(ดร. มหศักดิ์ เกตุฉ่ำ)



# # 5670385021 : MAJOR SOFTWARE ENGINEERING

KEYWORDS: CLOTHING RECOGNITION / CLOTHING CLASSIFICATION / COMPUTER VISION / BAG OF FEATURES

WISARUT SURAKARIN: Recognition Types of Clothing using Classification Method. ADVISOR: PROF. PRABHAS CHONGSTITVATANA, Ph.D., 68 pp.

The appearance of clothing represents the lifestyle, culture and social status. For example, we can recognize the people by the clothes they wear. Many contexts can be inferred from the type of clothes, for example, the people who wearing a suit shows that they were in a situation of the business. However, recognizing the type of clothes is a challenging problem in the area at computer vision. This thesis proposed a new approach to recognizing and classifying types of clothing by using the combination of SURF and LBP based on BoF. The input is an image of the type of fashion catalog where the clothes are fully exposed with models showing their faces.

For the preprocessing and features extraction of the Bag of Features is employed. There are four steps in the proposed recognition and classification method. Firstly, the cloth in an image is identified and located, then it is segmented by GrabCut. Secondly, the area in the image of cloth is divided into three sub-windows which are right-side, center and left-side. Thirdly, the feature extraction, SURF and LBP are applied to each sub-window to create a codebook. Finally, the classification is done by Support Vector Machine.

Our dataset consists of total 1131 images out of which the training set is 991 images and the remainder is the testing set. We separate types into seven categories of clothing image which included, jacket, shirt, suit, sweater, t-shirt, polo shirt and tank top. The result of the experiment illustrates that the proposed method can recognize types of clothing images accurately 73.57%.

Department: Computer Engineering      Student's Signature .....

Field of Study: Software Engineering      Advisor's Signature .....

Academic Year: 2014

## กิตติกรรมประกาศ

ข้าพเจ้าขอขอบพระคุณ ศาสตราจารย์ ดร. ประภาส จงสฤษดิ์วัฒนา อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่เสียสละเวลาช่วยเหลือ ให้คำปรึกษา คำแนะนำ แนวคิดและข้อคิดเห็นทั้งในเรื่องการเรียนรู้อันทางวิชาการตลอดจนการใช้ชีวิตในสังคม อีกทั้งให้ความเอาใจใส่อย่างสูง ทำให้ผู้วิจัยสามารถดำเนินงานวิจัยจนกระทั่งประสบผลสำเร็จลุล่วงไปด้วยดี ขอขอบพระคุณคณาจารย์ทุกท่านที่แนะนำ สั่งสอน และให้ความรู้แก่ข้าพเจ้าตลอดระยะเวลาการศึกษา

ขอขอบคุณคณะกรรมการสอบวิทยานิพนธ์ รองศาสตราจารย์ ดร. วิวัฒน์ วัฒนาวุฒิ และ ดร. มหศักดิ์ เกตุอำ่า กรรมการสอบวิทยานิพนธ์ ที่กรุณาสละเวลา ให้คำแนะนำ เทคนิคต่างๆ ทำให้วิทยานิพนธ์นี้มีความถูกต้องและสมบูรณ์มากขึ้น

ขอขอบคุณพี่ๆ เพื่อนๆ ญาติอาจารย์ที่ปรึกษาเดียวกัน ที่ช่วยเหลือและให้คำปรึกษา อภิปรายพูดคุยในการทำวิทยานิพนธ์ ทำให้ได้เห็นมุมมองต่างๆ ที่เป็นประโยชน์สำหรับการทำงานมากขึ้น อีกทั้งพี่ๆ เพื่อนๆ สาขาวิชาวิศวกรรมซอฟต์แวร์ สาขาวิศวกรรมวิทยาการคอมพิวเตอร์ และชมรมบัณฑิตศึกษา ที่ทำให้การใช้ชีวิตปริญญาโทมีสีสันและสนุกสนานตลอดระยะเวลาการศึกษา

ท้ายที่สุดขอขอบคุณสถานที่ต่างๆ ห้องปฏิบัติการวิจัยวิศวกรรมซอฟต์แวร์ ห้องปฏิบัติการระบบอัจฉริยะ หอสมุดกลาง และอาคารเฉลิมราชสุดาภิเษาสถาน สำหรับสถานที่ทำงาน ที่ทำกิจกรรมอันแสนอบอุ่นตลอดเวลาที่ทำงานร่วมกัน

## สารบัญ

หน้า

บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญตาราง.....	ฅ
สารบัญรูป.....	ฉ
บทที่ 1 บทนำ.....	1
1.1 ที่มาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของการวิจัย.....	2
1.3 ขอบเขตและข้อจำกัดของงานวิจัย.....	2
1.3.1 ขอบเขตของงานวิจัย.....	2
1.3.2 ข้อจำกัดของงานวิจัย.....	2
1.4 ลำดับขั้นตอนในการนำเสนองานวิจัย.....	3
1.5 ประโยชน์ที่คาดว่าจะได้รับ.....	3
1.6 ผลงานที่ตีพิมพ์จากวิทยานิพนธ์.....	4
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....	5
2.1 ทฤษฎีที่เกี่ยวข้อง.....	5
2.1.1 แบบจำลองของสี (Color Model).....	5
2.1.1.1 ปริภูมิสีแบบอาร์จีบี (RGB color space).....	5
2.1.1.2 ปริภูมิสีแบบเอชเอสวี (HSV color space).....	6
2.1.2 ฮิสโตแกรมของภาพ (Image Histogram).....	7
2.1.3 การปรับเท่าฮิสโตแกรม (Histogram Equalization).....	8

2.1.4 การคอนโวลูชัน (Convolution).....	8
2.1.5 Haar-like Feature.....	9
2.1.6 การค้นหาใบหน้าด้วยวิธีของ Viola และ Jones.....	10
2.1.6.1 เทคนิคการรวมภาพ (Integral Image) .....	11
2.1.6.2 ขั้นตอนการเรียนรู้ด้วยเอดาบูซ (AdaBoost).....	12
2.1.6.3 ขั้นตอนการจำแนกรูปภาพด้วย Cascade classifier .....	13
2.1.7 ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine).....	14
2.1.8 ชุดไลบรารีของ OpenCV (Open Source Computer Vision library).....	15
2.2 งานวิจัยที่เกี่ยวข้อง.....	16
บทที่ 3 การดึงคุณลักษณะสำคัญของภาพ.....	18
3.1 การอธิบายลักษณะพื้นผิวของภาพ.....	18
3.1.1 Local Binary Patterns (LBP).....	18
3.1.2 Local Directional Patterns (LDP) .....	19
3.2 การหาจุดสนใจของภาพ.....	21
3.2.1 การหาจุดสนใจด้วย SIFT (Scale Invariant Feature Transform).....	21
3.2.1.1 การเปลี่ยนแปลงระดับของภาพ.....	21
3.2.1.2 การกำหนดตำแหน่งจุดสนใจ.....	22
3.2.2 การสร้างเอกลักษณ์ของแต่ละจุดด้วย SIFT (Interest Point Description).....	23
3.2.2.1 การกำหนดทิศทางของจุดสนใจ (Orientation Assignment).....	23
3.2.2.2 การอธิบายลักษณะเด่นของภาพ (Keypoint Descriptor) .....	25
3.2.3 การหาจุดสนใจด้วย SURF (Speeded-Up Robust Features) .....	25
3.2.3.1 เทคนิคการรวมภาพ (Integral Image) .....	26
3.2.3.2 Hessian Detector.....	26



3.2.3.3 การเปลี่ยนแปลงระดับของภาพ.....	27
3.2.3.4 การกำหนดตำแหน่งจุดสนใจ.....	28
3.2.4 การสร้างเอกลักษณ์ของแต่ละจุดด้วย SURF (Interest Point Description).....	28
3.2.4.1 การกำหนดทิศทางของจุดสนใจ (Orientation Assignment).....	29
3.2.4.2 การอธิบายลักษณะเด่นของภาพ (KeyPoint Descriptor) .....	29
3.3 คุณลักษณะสำคัญครอบคลุม (Global Feature).....	30
3.3.1 Bag of Features (BoF) .....	30
3.3.1.1 การดึงคุณลักษณะสำคัญของจุดสนใจ .....	30
3.3.1.2 สร้างพจนานุกรมของจุดสนใจ .....	30
3.3.1.3 Histogram of Features.....	32
บทที่ 4 ขั้นตอนการดำเนินงานและการออกแบบ .....	33
4.1 การแบ่งส่วนของเสื้อผ้า (Clothing segmentation) .....	33
4.1.1 การค้นหาใบหน้าและร่างกายส่วนบน .....	35
4.1.2 การแบ่งส่วนภาพฉากหน้าและฉากหลัง .....	36
4.1.3 การแบ่งส่วนสีผิว .....	38
4.2 การดึงคุณลักษณะสำคัญของเสื้อผ้า (Features extraction).....	39
4.2.1 การจัดการข้อมูลนำเข้าก่อนการวิเคราะห์ (Data preprocessing).....	40
4.2.2 การดึงคุณลักษณะสำคัญของเสื้อผ้า (Clothing features extraction).....	41
4.3 การสอนให้เครื่องเรียนรู้และการทดสอบ (Clothing Classification).....	43
4.3.1 ขั้นตอนการเตรียมข้อมูลทดสอบ.....	45
4.3.2 วิธีการวัดประสิทธิผลความถูกต้อง .....	47
บทที่ 5 การทดลองและผลการวิจัย .....	49
5.1 ผลการทดสอบความถูกต้องในแต่ละวิธี .....	49

5.2 ผลการเปรียบเทียบกับวิธีอื่นๆ.....	59
บทที่ 6 สรุปผลการวิจัยและข้อเสนอแนะ .....	61
6.1 สรุปผลการวิจัย.....	61
6.2 ข้อเสนอแนะ .....	61
รายการอ้างอิง .....	63
ประวัติผู้เขียนวิทยานิพนธ์ .....	68



## สารบัญตาราง

	หน้า
ตารางที่ 2.1 แสดงจุดเด่น จุดด้อยของงานวิจัยที่เกี่ยวข้องพร้อมทั้งเทคนิควิธีที่ใช้ .....	16
ตารางที่ 4.1 จำนวนรูปภาพที่ใช้ในแต่ละกลุ่ม .....	45
ตารางที่ 4.2 การวัดประสิทธิผลของข้อมูล .....	47
ตารางที่ 4.3 สูตรการคำนวณค่า Accuracy, Precision และ Recall .....	47
ตารางที่ 4.4 k-fold cross-validation .....	48
ตารางที่ 5.1 Confusion matrix ของแต่ละชนิดเสื้อผ้าด้วยวิธี SIFT.....	50
ตารางที่ 5.2 รายละเอียดของค่าความแม่นยำ ค่าความครบถ้วนและค่า F-score ด้วยวิธี SIFT .....	50
ตารางที่ 5.3 Confusion matrix ของแต่ละชนิดเสื้อผ้าด้วยวิธี SURF.....	51
ตารางที่ 5.4 รายละเอียดของค่าความแม่นยำ ค่าความครบถ้วนและค่า F-score ด้วยวิธี SURF .....	52
ตารางที่ 5.5 Confusion matrix ของแต่ละชนิดเสื้อผ้าด้วยวิธี SIFT ร่วมกับ LBP.....	53
ตารางที่ 5.6 รายละเอียดของค่าความแม่นยำ ค่าความครบถ้วนและค่า F-score ด้วยวิธี SIFT ร่วมกับ LBP .....	54
ตารางที่ 5.7 Confusion matrix ของแต่ละชนิดเสื้อผ้าด้วยวิธี SURF ร่วมกับ LBP.....	55
ตารางที่ 5.8 รายละเอียดของค่าความแม่นยำ ค่าความครบถ้วนและค่า F-score ด้วยวิธี SURF ร่วมกับ LBP .....	56
ตารางที่ 5.9 Confusion matrix ของแต่ละชนิดเสื้อผ้าด้วยวิธีที่นำเสนอ.....	57
ตารางที่ 5.10 รายละเอียดของค่าความแม่นยำ ค่าความครบถ้วนและค่า F-score ด้วยวิธีที่ นำเสนอ.....	58
ตารางที่ 5.11 เปรียบเทียบวิธีต่างๆ ที่ได้ทดลองด้วยค่าความแม่นยำ ค่าความครบถ้วนและค่า F-score .....	59

## สารบัญรูป

	หน้า
รูปที่ 2.1 ปริภูมิสีแบบอาร์จีบี .....	6
รูปที่ 2.2 ปริภูมิสีแบบเอชเอสวี .....	6
รูปที่ 2.3 ฮีสโตแกรมของภาพ .....	8
รูปที่ 2.4 การปรับเท่าฮีสโตแกรมของภาพ .....	8
รูปที่ 2.5 ตัวอย่างการทำคอนโวลูชันในภาพ .....	9
รูปที่ 2.6 ลักษณะของ Haar Wavelet.....	10
รูปที่ 2.7 รูปแบบของ Features สำหรับตรวจจับลักษณะในแต่ละกลุ่ม .....	10
รูปที่ 2.8 ลักษณะเด่นทั้ง 3 แบบ คือ รูปสี่เหลี่ยม 2 รูป รูปสี่เหลี่ยม 3 รูป และรูปสี่เหลี่ยม 4 รูป ...	11
รูปที่ 2.9 ตัวอย่างการหาผลรวมของความเข้ม .....	12
รูปที่ 2.10 การทำงานของเอดาบูซ .....	13
รูปที่ 2.11 ขั้นตอนการจำแนกประเภทภาพใบหน้า .....	14
รูปที่ 2.12 ตัวอย่างการเรียนรู้ของเอดาบูซ .....	14
รูปที่ 2.13 ไฮเปอร์เพลนที่แบ่งชุดข้อมูลสองกลุ่ม .....	15
รูปที่ 3.1 วิธีการคำนวณ Local Binary Patterns .....	18
รูปที่ 3.2 ตัวอย่างรูปหลังจากการคำนวณด้วย Local Binary Patterns .....	19
รูปที่ 3.3 Filter mask ทั้ง 8 ทิศทาง.....	19
รูปที่ 3.4 ภาพหลังผ่าน Filter mask แต่ละทิศทาง .....	20
รูปที่ 3.5 ผลการจัดอันดับและแปลงเป็นเลขฐานสิบของจุดภาพ.....	20
รูปที่ 3.6 ผลของภาพที่ได้จากการคำนวณ LBP .....	20
รูปที่ 3.7 ภาพที่ได้จากการลดขนาด 4 ระดับและรายละเอียด 5 ระดับ .....	21
รูปที่ 3.8 การทำ Difference of Gaussian .....	22

รูปที่ 3.9 การหาค่ามากที่สุดและน้อยสุดของจุด X กับจุดรอบข้างทั้ง 26 จุด .....	23
รูปที่ 3.10 ตัวอย่างจุดสนใจ .....	24
รูปที่ 3.11 ค่าฮีสโตแกรมทั้ง 36 ช่วง .....	24
รูปที่ 3.12 จุดรอบๆ จุดสนใจ .....	25
รูปที่ 3.13 การเปลี่ยนค่าของทิศทางเป็นฮีสโตแกรม .....	25
รูปที่ 3.14 ตัวอย่างการคำนวณพื้นที่ที่ต้องการ .....	26
รูปที่ 3.15 ตัวอย่างตัวกรองในทิศทางของ x, y และ xy .....	27
รูปที่ 3.16 การสเกลขนาดของตัวกรองที่มีขนาดต่างกัน .....	27
รูปที่ 3.17 ตัวอย่างสเกลในแต่ละระดับขั้น .....	27
รูปที่ 3.18 การเปรียบเทียบกับพิกเซลข้างเคียง .....	28
รูปที่ 3.19 ตัวอย่าง Haar wavelet ในทิศทาง x และทิศทาง y .....	28
รูปที่ 3.20 ทิศทางของตัวกรอง Haar wavelet .....	29
รูปที่ 3.21 ผลรวมของ Haar wavelet ของแต่ละพื้นที่ย่อย .....	29
รูปที่ 3.22 คุณภาพของคุณลักษณะสำคัญ .....	30
รูปที่ 3.23 ตัวอย่างพจนานุกรมของภาพ .....	32
รูปที่ 3.24 ตัวอย่าง Histogram of Features .....	32
รูปที่ 3.25 สัดส่วนของการเขียนแบบร่างมนุษย์ .....	35
รูปที่ 5.1 กราฟแสดงค่าของความแม่นยำและความถูกต้องในแต่ละกลุ่มการทดสอบด้วยวิธี SIFT ...	51
รูปที่ 5.2 กราฟแสดงค่าของความแม่นยำและความถูกต้องในแต่ละกลุ่มการทดสอบด้วยวิธี SURF .	53
รูปที่ 5.3 กราฟแสดงค่าของความแม่นยำและความถูกต้องในแต่ละกลุ่มการทดสอบด้วยวิธี SIFT ร่วมกับ LBP .....	55
รูปที่ 5.4 กราฟแสดงค่าของความแม่นยำและความถูกต้องในแต่ละกลุ่มการทดสอบด้วยวิธี SURF ร่วมกับ LDP .....	57

รูปที่ 5.5 กราฟแสดงค่าของความแม่นยำและความถูกต้องในแต่ละกลุ่มการทดสอบด้วยวิธีที่ นำเสนอ.....	59
รูปที่ 5.6 กราฟการเปรียบเทียบในแต่ละวิธีที่ได้ทดลอง โดยแบ่งผลความถูกต้องตามชนิดของ เสื้อผ้า.....	60



# บทที่ 1

## บทนำ

### 1.1 ที่มาและความสำคัญของปัญหา

ในอดีตมนุษย์สวมใส่เสื้อผ้าเพื่อปกป้องร่างกายของตนไม่ให้เกิดความอันตรายจากสิ่งแวดล้อม เช่น สิ่งแวดล้อมทางอากาศ ความร้อนจากแสงแดด ความเย็น ฝน และความอันตรายอื่นๆ เป็นต้น ซึ่งการออกแบบเสื้อผ้าก็จะทำให้มีความแตกต่างกันตามความเหมาะสมของแต่ละสิ่งแวดล้อมและเปลี่ยนไปตามยุคสมัย ซึ่งในปัจจุบันเสื้อผ้าทำหน้าที่มากกว่าเป็นเครื่องนุ่งห่มหรือเพื่อปกป้องร่างกายเพียงอย่างเดียว แต่เสื้อผ้าจะเป็นสิ่งที่สามารถสะท้อนถึงสถานะทางสังคม วิธีการดำเนินชีวิต และโดยเฉพาะอย่างยิ่ง คือ บ่งบอกถึงวัฒนธรรมที่แตกต่างกันของแต่ละพื้นที่ ดังนั้นเสื้อผ้าจึงมีความสำคัญที่ใช้สำหรับอธิบายบุคคลอื่นๆ ในสังคม ตัวอย่างเช่น “ผู้คนสวมใส่เสื้อสูทสีดำ” จะสามารถตีความหมายได้ว่าพวกเขากำลังอยู่ในสถานการณ์ของการทำธุรกิจและอาจจะหมายถึงงานที่เป็นทางการ หรือ “ผู้คนสวมใส่กางเกงว้ายน้ำ” จะแสดงให้เห็นว่าพวกเขากำลังออกกำลังกายทางน้ำ เป็นต้น ตัวอย่างอื่นๆ ที่ยังสามารถใช้ในการอธิบายบุคคล เช่น การสวมใส่เสื้อแขนยาวจะหมายถึงอยู่ในสภาวะอากาศที่เย็นหรือหนาว การสวมใส่เสื้อแขนสั้นหรือเสื้อกล้ามจะหมายถึงอยู่ในช่วงฤดูร้อน เป็นต้น

เสื้อผ้าของแต่ละบุคคลมีประโยชน์ในการช่วยจดจำบุคคลนั้นได้ ซึ่งเมื่อเปรียบเทียบกับการรู้จำใบหน้าของคนแล้ว เสื้อผ้าสามารถช่วยให้รู้จำบุคคลนั้นได้เร็วยิ่งขึ้นเพราะเราเห็นเสื้อผ้าในระยะไกลก็สามารถบอกได้ว่าเป็นใคร ซึ่งชนิดของเสื้อผ้าสามารถบอกรายละเอียดของบุคคลนั้นให้มีส่วนในการรู้จำได้ดียิ่งขึ้น ดังนั้นงานวิจัยฉบับนี้จะนำเสนอวิธีรู้จำชนิดของเสื้อผ้า โดยการจำแนกประเภท

อย่างไรก็ตามการรู้จำชนิดของเสื้อผ้าเป็นปัญหาที่มีความท้าทายมากในด้านของการประมวลผลภาพและการเรียนรู้ของเครื่องเพราะเสื้อผ้าเปลี่ยนแปลงรูปร่างหรือท่าทางอยู่ตลอดเวลา และผู้คนที่สวมใส่เสื้อผ้าจะมีลักษณะที่ต่างกัน ชนิดของเสื้อผ้าที่มีลักษณะคล้ายคลึงกัน โดยในงานวิจัยนี้ได้นำทฤษฎีของการสร้างคุณลักษณะสำคัญครอบคลุมมาใช้ในการอธิบายและค้นหาจุดสำคัญของเสื้อผ้าในแต่ละชนิดเพื่อแก้ไขปัญหาของการเปลี่ยนแปลงรูปร่างหรือท่าทางของเสื้อผ้า

วิทยานิพนธ์ฉบับนี้ได้ใช้ทฤษฎีทางการประมวลผลภาพมาประยุกต์ใช้เพื่อการรู้จำชนิดของเสื้อผ้า โดยงานวิจัยนี้ได้เลือกใช้ชุดไลบรารีของ OpenCV และได้เลือกใช้ภาษา C++ ของชุดเครื่องมือ Qt creator เป็นหลัก

## 1.2 วัตถุประสงค์ของการวิจัย

- 1) เพื่อรู้จำชนิดของเสื้อผ้าได้ถูกต้องแม่นยำมากขึ้นในค่าที่ยอมรับได้
- 2) เพื่อปรับปรุงการรู้จำชนิดของเสื้อผ้าในขั้นตอนของการดำเนินการเรียนรู้ของเครื่องและการทดสอบ
- 3) เพื่อสร้างวิธีการใหม่ โดยได้นำหลักการของคุณลักษณะสำคัญท้องถิ่นและคุณลักษณะสำคัญครอบคลุมมาใช้

## 1.3 ขอบเขตและข้อจำกัดของงานวิจัย

### 1.3.1 ขอบเขตของงานวิจัย

- 1) สามารถตรวจจับตำแหน่งของเสื้อผ้าในรูปภาพ แต่จะไม่จดจำข้อมูลไว้ในระบบ และเสื้อผ้าที่ตรวจจับนั้นจะใช้เพียงเสื้อผ้าส่วนบนเท่านั้น
- 2) สามารถแบ่งส่วนข้อมูลภาพที่มีเสื้อผ้า โดยแบ่งส่วนระหว่างเสื้อผ้ากับส่วนพื้นหลังหรือส่วนที่ไม่เกี่ยวข้องได้ถูกต้องแม่นยำมากขึ้น
- 3) สามารถวิเคราะห์และสร้างคุณลักษณะสำคัญจากเสื้อผ้าในภาพที่มีเสื้อผ้าเพื่อใช้ในการทดสอบระบบได้
- 4) สามารถรู้จำภาพเสื้อผ้าได้ว่าเป็นเสื้อผ้าชนิดอะไร โดยในวิทยานิพนธ์ฉบับนี้ได้กำหนดกลุ่มของชนิดเสื้อผ้าไว้ทั้งหมด 7 กลุ่ม คือ แจ็คเก็ต เสื้อเชิ้ต เสื้อยืด เสื้อโปโล เสื้อสูท เสื้อกันหนาว และเสื้อกล้าม
- 5) กลุ่มของเสื้อผ้าทั้ง 7 กลุ่มนั้นในที่นี้ผู้วิจัยไม่ได้ทำการแบ่งแยกเพศตามแต่ละชนิดเสื้อผ้า

### 1.3.2 ข้อจำกัดของงานวิจัย

- 1) รูปภาพหรือภาพจากวิดีโอที่นำมาใช้ในโปรแกรมต้นแบบต้องเป็นภาพหน้าตรงเท่านั้น ไม่มีการปกปิดบางส่วน เช่น หน้า ตา ปาก และเสื้อผ้า เป็นต้น
- 2) ภาพของพื้นหลังจะต้องไม่มีสิ่งกีดขวางหรือมีสีสันทึบที่คล้ายกับสีเสื้อผ้า โดยในที่นี้ได้เลือกใช้พื้นหลังขาวในการทดสอบ
- 3) ภาพของบุคคลที่เดินผ่านหน้ากล้องต้องอยู่ในระยะที่เหมาะสมไม่ใกล้หรือไกลจนเกินไป และภาพต้องไม่เบลอจนไม่สามารถตรวจจับได้ว่าเป็นภาพบุคคลหรือไม่
- 4) ชนิดของเสื้อผ้าที่ใช้ทดสอบและรู้จำ ต้องเป็นชนิดของเสื้อผ้าที่อยู่ในรูปแบบที่กำหนดไว้ใน 7 กลุ่มเท่านั้น



- 5) เสื้อผ้าที่ใช้ในการทดสอบระบบจะต้องเป็นเสื้อผ้าเพียงชั้นเดียวหรือหนึ่งชนิดเท่านั้น ระบบไม่สามารถตรวจจับเสื้อผ้าที่ใส่ซ้อนทับกัน ซึ่งอาจจะทำให้โปรแกรมเกิดความผิดพลาดขึ้นได้

#### 1.4 ลำดับขั้นตอนในการนำเสนองานวิจัย

- 1) ศึกษาการประมวลผลภาพเบื้องต้นและวิธีการเรียนรู้ของเครื่อง
- 2) ศึกษาวิธีการค้นหาของเสื้อผ้าในภาพและระบุตำแหน่งของเสื้อผ้าได้ถูกต้อง
- 3) ศึกษาขั้นตอนการแบ่งส่วนข้อมูลภาพเพื่อจัดส่วนที่ไม่ใช่เสื้อผ้าออก ซึ่งจะต้องรวมทั้งสีผิวของคนด้วย
- 4) ศึกษาและกำหนดชนิดของเสื้อผ้าที่จะนำมาใช้ในการรู้จำว่าต้องมีกี่ชนิด ชนิดอะไรบ้าง และแต่ละชนิดแตกต่างจากชนิดอื่นๆ อย่างไร
- 5) ศึกษา วิเคราะห์ และสร้างคุณลักษณะสำคัญจากความสัมพันธ์ของข้อมูลภายในภาพเสื้อผ้าว่าจะใช้วิธีการอะไรบ้าง
- 6) ทดสอบผลการรู้จำว่าอยู่ในค่าความเชื่อมั่นหรือไม่
- 7) ออกแบบโครงสร้าง หน้าจอและเครื่องมือที่ใช้ของวิธีการจำแนกชนิดของเสื้อผ้า
- 8) สร้างระบบต้นแบบจากโครงสร้างที่ได้กำหนดเอาไว้
- 9) สรุปผลการวิจัยและข้อเสนอแนะ
- 10) จัดทำรายงานวิทยานิพนธ์

#### 1.5 ประโยชน์ที่คาดว่าจะได้รับ

- 1) ช่วยเพิ่มประสิทธิภาพในการรู้จำชนิดของเสื้อผ้าให้มีความถูกต้องมากขึ้นและสามารถนำไปปรับใช้กับการรู้จำเสื้อผ้าชนิดอื่นๆ มากขึ้นได้
- 2) ช่วยลดระยะเวลาในการรู้จำชนิดของเสื้อผ้าในขั้นตอนการหาคุณลักษณะสำคัญท้องถิ่นและคุณลักษณะสำคัญครอบคลุม

## 1.6 ผลงานที่ตีพิมพ์จากวิทยานิพนธ์

งานวิจัยฉบับนี้ได้รับการตีพิมพ์เป็นบทความเชิงวิชาการเรื่อง “Predicting Types of Clothing Using SURF and LDP based on Bag of Features” โดย วิศรุต สุรการินทร์ และ ประภาส จงสถิตย์วัฒนา ได้รับคัดเลือกและตีพิมพ์ในงานประชุมวิชาการ “2015 12<sup>th</sup> International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON 2015)” ระหว่างวันที่ 24-27 มิถุนายน 2558 ณ โรงแรมโนโวเทล หัวหิน ชะอำบีช รีสอร์ท แอนด์ สปา ประเทศไทย



## บทที่ 2

### ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในบทนี้จะแบ่งออกเป็น 2 ส่วน คือ ส่วนแรกจะกล่าวถึงทฤษฎีที่เกี่ยวข้องในงานวิจัยนี้ ซึ่งจะเกี่ยวกับการประมวลผลภาพและวิธีการรู้จำชนิดของเสื้อผ้า และส่วนที่สองจะกล่าวถึงงานวิจัยก่อนหน้าที่มีคนเคยทำเรื่องการวิเคราะห์เสื้อผ้ามาก่อน โดยในส่วนนี้จะระบุถึงรายละเอียด ข้อดี ข้อเสียของแต่ละงานวิจัย

#### 2.1 ทฤษฎีที่เกี่ยวข้อง

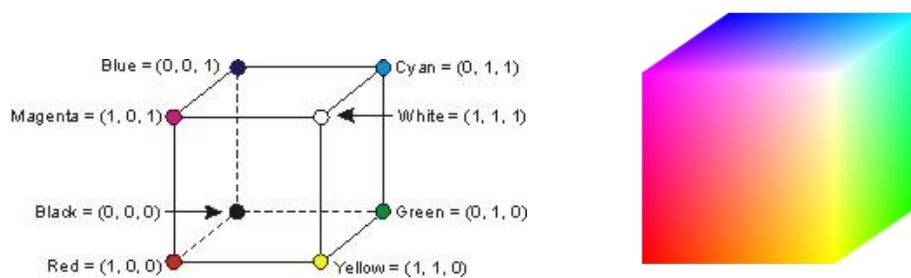
ในการวิเคราะห์ข้อมูลภายในภาพจำเป็นต้องใช้ทฤษฎี ในการประมวลผลภาพและการเรียนรู้ของเครื่องเป็นหลัก ซึ่งผู้วิจัยได้ศึกษาทฤษฎีต่างๆ ที่เกี่ยวข้อง อันได้แก่ การกำจัดสัญญาณรบกวน การเตรียมข้อมูล การแบ่งส่วนข้อมูลภาพ การค้นหา วิเคราะห์และดึงคุณลักษณะสำคัญภายในภาพ เพื่อใช้ในการเรียนรู้ของเครื่อง โดยมีรายละเอียดดังต่อไปนี้

##### 2.1.1 แบบจำลองของสี (Color Model)

มาตรฐานของสีที่ใช้อยู่ในปัจจุบันมีอยู่หลากหลายแบบจำลองด้วยกันขึ้นอยู่กับนำไปใช้ ซึ่งในด้านของการประมวลผลภาพและวีดิทัศน์ ค่าของข้อมูลสีจากภาพดิจิทัลนั้นจะถูกเก็บอยู่ในรูปแบบขององค์ประกอบฐานสี (Color basis) เข้าเป็นสีต่างๆ ในหนึ่งจุดภาพ (Pixel) ดังนั้นจึงสามารถนำไปอ้างอิงกับปริภูมิสี (Color space) ได้หลากหลายรูปแบบ ตัวอย่างเช่น ปริภูมิสีแบบอาร์จีบี (RGB color space) ที่ถูกนำไปใช้อย่างกว้างขวางในสื่อดิจิทัลและการตกแต่งภาพและอุปกรณ์อิเล็กทรอนิกส์ทั่วไป เป็นต้น โดยในวิทยานิพนธ์จะใช้แบบจำลองของสี ดังนี้

##### 2.1.1.1 ปริภูมิสีแบบอาร์จีบี (RGB color space)

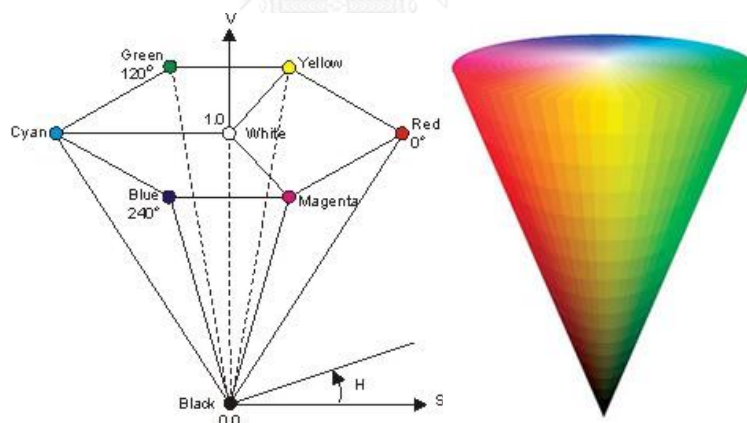
ปริภูมิสีแบบอาร์จีบี [1] เป็นปริภูมิที่สามารถพบเห็นได้ทั่วไปในอุปกรณ์อิเล็กทรอนิกส์ เช่น กล้องถ่ายรูป โทรศัพท์มือถือ และจอโทรทัศน์ เป็นต้น ซึ่งแนวคิดของปริภูมิสีแบบอาร์จีบีได้กล่าวไว้ว่า สีใดๆ ก็ตามเกิดจากการผสมแม่สี 3 สี คือ แดง เขียว และน้ำเงิน ที่สอดคล้องกับเซลล์รับแสงของตามนุษย์ คือ ในย่านของ แดง เขียว และน้ำเงิน สามารถแทน ด้วยแบบจำลองทรงลูกบาศก์ ดังรูปที่ 2.1



รูปที่ 2.1 ปริภูมิสีแบบอาร์จีบี

### 2.1.1.2 ปริภูมิสีแบบเอชเอสวี (HSV color space)

ปริภูมิสีแบบเอชเอสวี เป็นการพิจารณาสีโดยใช้องค์ประกอบของสี คือ เนื้อสี (Hue) ของสีแดง สีเขียว และสีน้ำเงิน ซึ่งถ้าเนื้อสีมีค่าเท่ากับศูนย์จะเป็นสีแดงแต่เมื่อค่าของเนื้อสีมีค่าเพิ่มขึ้นเรื่อยๆ จนถึง 255 ก็จะมาวนกลับมาเป็นสีแดงใหม่อีกครั้ง ดังนั้นแม่สีสามารถแทนให้อยู่ในรูปขององศาได้ คือ สีแดงเท่ากับ 0 องศา สีเขียวเท่ากับ 120 องศา และสีน้ำเงิน 240 องศา ความสดของสี (Saturation) คือ ค่าแสงผสมกับค่าเนื้อสี ซึ่งถ้ามีค่าเท่ากับ 0 แล้วสีที่ได้จะเป็นสีขาว แต่ถ้ามีค่าเท่ากับ 255 แสดงว่าจะไม่มีสีขาวอยู่เลย ส่วนค่าน้ำหนักของสี (Value) คือความสว่างของสีซึ่งสามารถวัดได้โดยค่าความเข้มของความสว่างในแต่ละสีที่ประกอบกัน ดังรูปที่ 2.2



รูปที่ 2.2 ปริภูมิสีแบบเอชเอสวี [2]

การแปลงภาพสีในแบบจำลองสีอาร์จีบีเป็นภาพจำลองสีเอชเอสวีนั้นจะมีวิธีการเปลี่ยนตามสมการที่ (1) โดยในสมการนั้นค่าของ MAX และ MIN คือ ค่าที่สูงที่สุดและค่าที่น้อยที่สุดของสีแดง สีเขียวและน้ำเงิน

$$H = \begin{cases} \left(\frac{G - B}{MAX - MIN}\right) \times 60 & \text{if } R = MAX \\ \left(2 + \frac{B - R}{MAX - MIN}\right) \times 60 & \text{if } G = MAX \\ \left(4 + \frac{R - G}{MAX - MIN}\right) \times 60 & \text{if } B = MAX \end{cases} \quad (1)$$

$$S = \frac{MAX - MIN}{MAX}$$

$$V = \frac{MAX}{255}$$

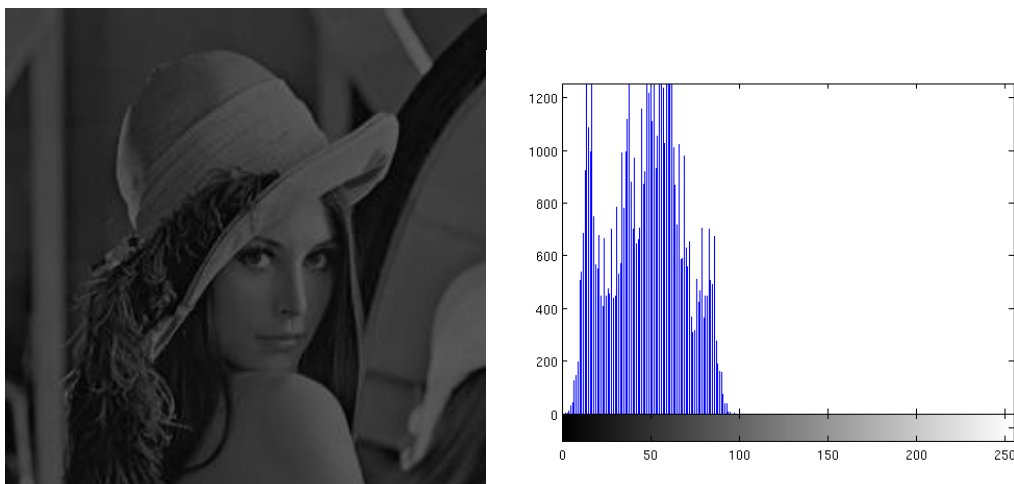
โดยที่ MAX คือค่าที่สูงที่สุดของสีแดง สีเขียวและน้ำเงิน  
MIN คือค่าที่น้อยที่สุดของสีแดง สีเขียวและน้ำเงิน

จากสมการที่ (1) จะเห็นได้ว่าแบบจำลองสีอาร์จีบีจะถูกปรับค่าใหม่ให้อยู่ในช่วง 0 ถึง 1 ซึ่งจะต้องทำการปรับค่าของเอชใหม่ให้อยู่ในช่วงของ 0 ถึง 360 ตามเงื่อนไขสมการที่ (2)

$$H = H + 360 \quad \text{if } H < 0 \quad (2)$$

### 2.1.2 ฮิสโตแกรมของภาพ (Image Histogram)

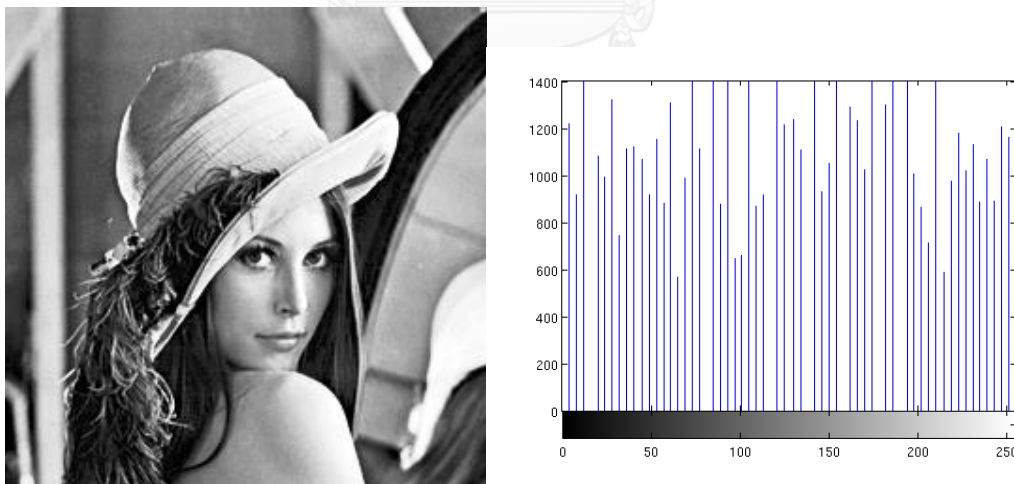
ฮิสโตแกรม [3] คือ มาตรการที่ใช้ในการบ่งบอกการกระจายตัวของความเข้มแต่ละจุดสีในภาพของภาพสีหรือภาพเฉดเทา ซึ่งการคำนวณค่าฮิสโตแกรมนี้ผลที่ได้จะอยู่ในรูปของกราฟแท่ง แสดงผลรวมความเข้มสีในแต่ละช่วงของภาพ โดยที่แนวนอน คือ ค่าระดับความเข้ม และแกนตั้ง คือ ความถี่ของความเข้มสีนั้นๆ ซึ่งฮิสโตแกรมจะสามารถบอกช่วงความเข้มของค่าสีออกมาเป็นกราฟว่าแต่ละภาพมีลักษณะกราฟเอียงหรือเกาะกลุ่มไปทางทิศที่สว่างหรือมืดเกินไป เช่น ในรูปที่ 2.3 จะเห็นได้ว่ากราฟมีลักษณะเกาะกลุ่มในช่วงค่า 0 ถึง 100 แสดงว่าภาพนี้มีลักษณะของภาพที่มืดเกินไป ซึ่งกราฟฮิสโตแกรมที่ดีควรมีการกระจายค่าเท่าๆ กันในทุกช่วงตั้งแต่ 0 ถึง 255



รูปที่ 2.3 ฮิสโตแกรมของภาพ [3]

### 2.1.3 การปรับเท่าฮิสโตแกรม (Histogram Equalization)

การปรับเท่าฮิสโตแกรม [4] เป็นการปรับปรุงให้กราฟมีลักษณะการกระจายของความเข้มสีแบบสม่ำเสมอตั้งแต่ 0 ถึง 255 ซึ่งผลจากการทำการปรับเท่าฮิสโตแกรมจะช่วยให้ภาพชัดเจนและคมชัดยิ่งขึ้นในภาพที่มีความแตกต่างกันน้อยของระดับสีที่มืดที่สุดในภาพกับที่สว่างที่สุดในภาพ (Low contrast) หรือภาพที่มีสีซีดจาง ดังรูปที่ 2.4 หลังใช้การปรับเท่าฮิสโตแกรมจะพบได้ว่าการกระจายตัวของกราฟมีความสม่ำเสมอ

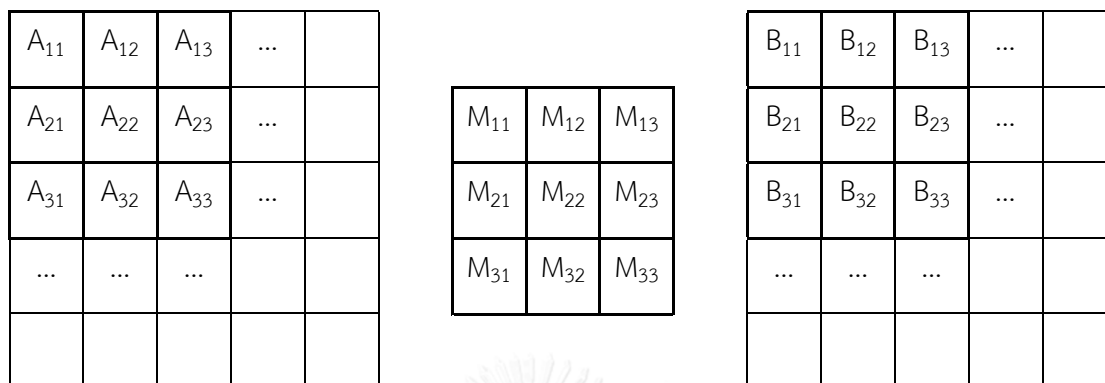


รูปที่ 2.4 การปรับเท่าฮิสโตแกรมของภาพ [4]

### 2.1.4 การคอนโวลูชัน (Convolution)

การคอนโวลูชันในภาพ คือ การคำนวณพื้นที่ที่รอบๆ ค่าของพิกเซล โดยการนำค่าของพิกเซลที่ล้อมรอบพิกเซลที่กำลังพิจารณามาทำการคำนวณกับชุดตัวเลขคงที่กลุ่มหนึ่งที่เรียกว่า ค่าตัวร่วม หรือ

Mask Coefficient เข้าไปคูณกับค่าของพิกเซลเหล่านั้น แล้วนำผลคูณที่ได้มาบวกเข้าด้วยกัน ผลลัพธ์จะเก็บไว้ในตำแหน่งของพิกเซลที่กำลังพิจารณาของภาพ ซึ่งหลักการทำคอนโวลูชันในภาพจะทำการสแกน หรือ sliding window จากตำแหน่งบนซ้ายไปยังตำแหน่งบนขวา แสดงดังสมการที่ (3)



รูปที่ 2.5 ตัวอย่างการทำคอนโวลูชันในภาพ

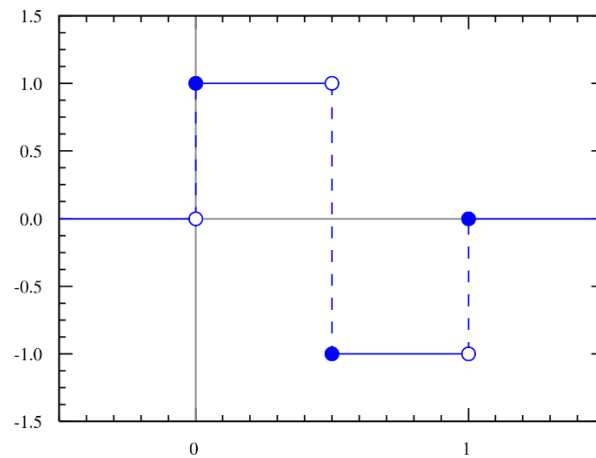
$$B(x, y) = \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} A(i, j) \cdot M(x - i, y - j) \quad (3)$$

ตัวอย่างการคำนวณคอนโวลูชันในภาพที่ตำแหน่ง  $x = 2$  และ  $y = 2$  ดังนี้

$$\begin{aligned} \text{Pixel}(2, 2) &= (A_{11} \times M_{11}) + (A_{12} \times M_{12}) + (A_{13} \times M_{13}) \\ &\quad + (A_{21} \times M_{21}) + (A_{22} \times M_{22}) + (A_{23} \times M_{23}) \\ &\quad + (A_{31} \times M_{31}) + (A_{32} \times M_{32}) + (A_{33} \times M_{33}) \end{aligned}$$

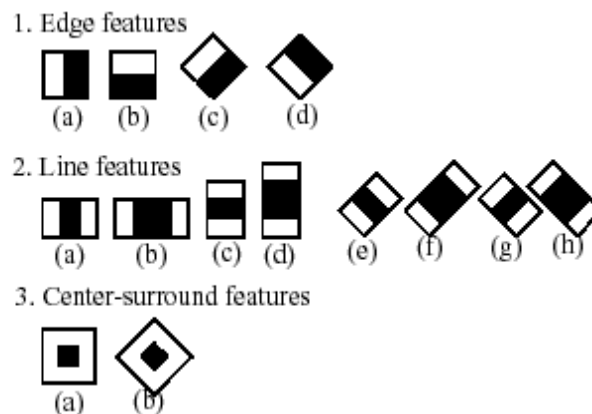
### 2.1.5 Haar-like Feature

รูปสี่เหลี่ยมมุมฉากที่เรียกว่า Haar Feature หรือ Haar-like Feature จะมีลักษณะคล้ายกับฟังก์ชันของ Haar Wavelets ซึ่งเป็น Wavelets พื้นฐานที่มีลักษณะเป็นฟังก์ชันไม่ต่อเนื่อง (Discrete Function) แต่ Wavelets ในลักษณะนี้จะมีข้อด้อย คือ ไม่มีความต่อเนื่องของฟังก์ชัน ดังรูปที่ 2.6



รูปที่ 2.6 ลักษณะของ Haar Wavelet [5]

ส่วน Haar-like Feature ตามวิธีของ Viola และ Jones [6] เป็นวิธีตรวจจับและตีความวัตถุภายในภาพด้วยการอาศัยชุดโครงสร้างรูปสี่เหลี่ยมที่ได้จากการฝึกสอนของข้อมูลตัวอย่าง ทั้งตัวอย่างที่ถูกต้องหรือตัวอย่างบวกและตัวอย่างที่ไม่ถูกต้องหรือตัวอย่างลบ ดังรูปที่ 2.7 โครงสร้างสี่เหลี่ยมที่ใช้ในการตรวจหาลักษณะต่างๆ โดยที่ค่าเฉลี่ยผลต่างระหว่างพื้นที่ส่วนสีขาวและพื้นที่ส่วนสีดำจะต้องมีค่ามากกว่าเส้นขีดแบ่ง (Threshold) ที่กำหนดในขั้นตอนการเรียนรู้ แบ่งออกเป็น 3 กลุ่ม คือ คุณลักษณะสำหรับตรวจจับเส้น คุณลักษณะสำหรับตรวจจับขอบและคุณลักษณะสำหรับตรวจจับบริเวณรอบจุดศูนย์กลาง ซึ่งคุณลักษณะที่ได้มายังสามารถเปลี่ยนแปลงขนาดและตำแหน่งได้



รูปที่ 2.7 รูปแบบของ Features สำหรับตรวจจับลักษณะในแต่ละกลุ่ม [6]

### 2.1.6 การค้นหาใบหน้าด้วยวิธีของ Viola และ Jones

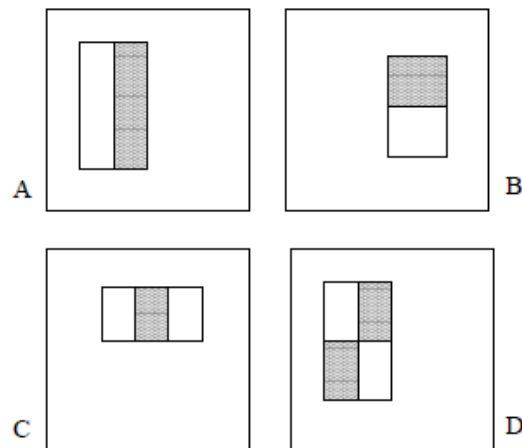
วิธีค้นหาใบหน้าของ Viola และ Jones [6] นั้นได้ถูกนำเสนอในปี 2001 ซึ่งเป็นวิธีการที่จัดอยู่ในกลุ่มของ Appearance-based methods เป็นโมเดลที่ได้จากการเรียนรู้ของกลุ่มรูปภาพตัวอย่างฝึกสอน ทั้งที่เป็นภาพใบหน้าหรือตัวอย่างบวกและไม่ใช่ภาพใบหน้าหรือตัวอย่างลบ โดย



วิเคราะห์ค่าทางสถิติและการเรียนรู้ของเครื่อง (Machine Learning) ซึ่งในปัจจุบันวิธีการของ Viola และ Jones นั้นได้รับการยอมรับและมีการใช้งานอย่างแพร่หลายเพราะมีความสามารถในการตรวจจับใบหน้าได้อย่างถูกต้อง รวดเร็วและมีประสิทธิภาพ ซึ่งสามารถนำไปใช้ในงานได้จริงทั้งในโปรแกรมตกแต่งรูปภาพและอุปกรณ์อิเล็กทรอนิกส์ เช่น กล้องถ่ายรูปดิจิทัล โทรศัพท์มือถือ เป็นต้น Viola และ Jones นั้นได้นำเสนอวิธีการค้นหาใบหน้าในภาพ โดยแบ่งเป็น 3 ขั้นตอนได้แก่ เทคนิคการรวมภาพ (Integral image) เอดาบูช (AdaBoost) และการจำแนกรูปภาพด้วย Cascade โดยในแต่ละขั้นตอนมีรายละเอียด ดังนี้

### 2.1.6.1 เทคนิคการรวมภาพ (Integral Image)

Viola และ Jones ได้นำการจำลองรูปแบบ Haar-like Feature ดังที่ได้กล่าวมาข้างต้น เป็นตัวตรวจจับใบหน้าในแต่ละลักษณะเด่นทั้งหมด 3 รูปแบบ คือ รูปสี่เหลี่ยมสองรูป รูปสี่เหลี่ยมสามรูป และรูปสี่เหลี่ยมสี่รูป ดังรูปที่ 2.8 ซึ่งในแต่ละลักษณะเด่นจะประกอบไปด้วยพื้นที่สี่เหลี่ยม 2 ประเภท คือ ส่วนที่เป็นสีขาว และส่วนที่เป็นสีเทา โดยที่ค่าเฉลี่ยผลต่างระหว่างพื้นที่ส่วนสีขาวและพื้นที่ส่วนสีดำจะต้องมีค่ามากกว่าเส้นขีดแบ่ง (Threshold) ที่ได้กำหนดในระหว่างการเรียนรู้ของขั้นตอนเอดาบูช

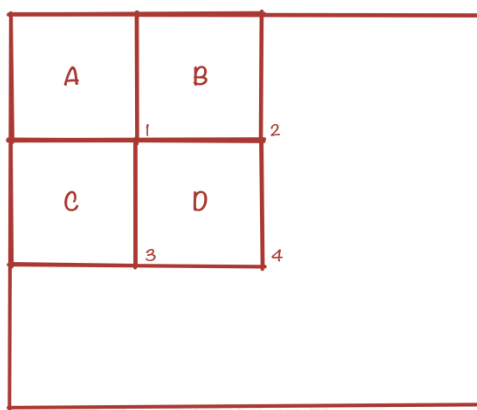


รูปที่ 2.8 ลักษณะเด่นทั้ง 3 แบบ คือ รูปสี่เหลี่ยม 2 รูป รูปสี่เหลี่ยม 3 รูป และรูปสี่เหลี่ยม 4 รูป

ซึ่งในการคำนวณหาสี่เหลี่ยมหรือ Haar-like Feature ได้อย่างรวดเร็วนั้น Viola และ Jones ได้นำเสนอเทคนิคที่เรียกว่า Integral Image โดยทั่วไปแล้ว Integrating ก็คือการรวมของหน่วยที่มีขนาดเล็กเข้าด้วยกัน ซึ่งในกรณีนี้หน่วยที่มีขนาดเล็กก็คือพิกเซลนั่นเอง การรวมค่าแต่ละพิกเซลก็คือผลรวมของพิกเซลทั้งหมดที่อยู่เหนือกว่าพิกเซลนั้น โดยเริ่มที่ซ้ายบนสุดของภาพแล้วลากไปตามขวางทางขวาและลงล่าง ขั้นตอนการคำนวณความเข้มรวมของแต่ละพิกเซล โดยใช้สมการที่ (4) โดยที่  $I(x, y)$  คือ ค่าของเทคนิคการรวมภาพ และ  $II(x, y)$  คือ ภาพต้นแบบ

$$I(x, y) = \sum_{x' \leq x, y' \leq y} I(x', y') \quad (4)$$

ดังนั้นสามารถที่จะหา Integral Image ได้ทุกตำแหน่งในพื้นที่ของภาพ เช่นตัวอย่างต่อไปนี้ จากรูปที่ 2.9 กำหนดให้รูปสี่เหลี่ยมใหญ่เป็น Sub-window และให้สี่เหลี่ยม D เป็น Haar-like Feature ชนิดหนึ่งที่อยู่ใน Sub-window นั้น ถ้าต้องการหา Integral Image ในสี่เหลี่ยม D สามารถหาได้ดังนี้



รูปที่ 2.9 ตัวอย่างการหาผลรวมของความเข้ม

จากรูปที่ 2.9 เมื่อตำแหน่งที่หนึ่ง คือ ผลรวมของพิกเซลในสี่เหลี่ยม [A] ตำแหน่งที่สอง คือ ผลรวมของพิกเซลในสี่เหลี่ยม [A+B] ตำแหน่งที่สาม คือ ผลรวมของพิกเซลในสี่เหลี่ยม [A+C] และตำแหน่งที่สี่ คือ ผลรวมของพิกเซลในสี่เหลี่ยม [A+B+C+D] ดังนั้นจะได้ผลรวมในสี่เหลี่ยม [D] คือ  $[A+B+C+D] + [A] - [A+B] - [A+C]$  หรือเขียนได้ดังสมการที่ (5)

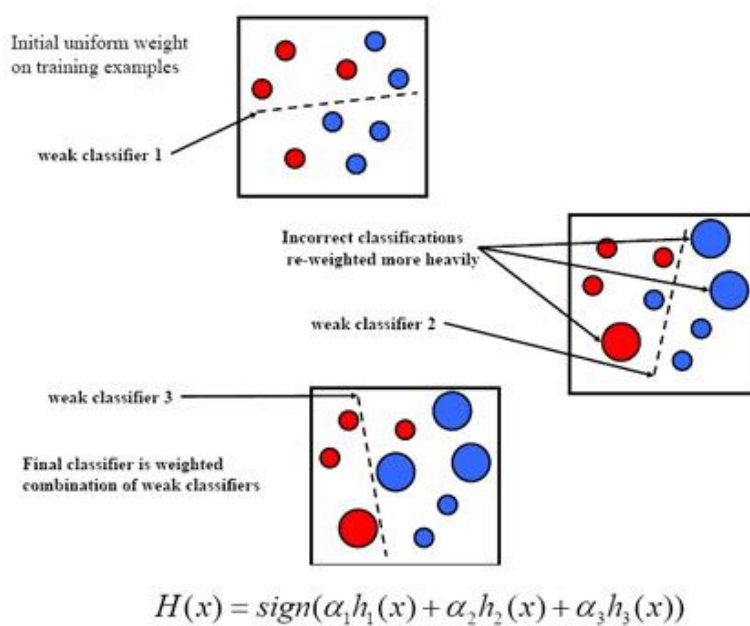
$$[D] = 4 + 1 - 2 - 3 \quad (5)$$

### 2.1.6.2 ขั้นตอนการเรียนรู้ด้วยเอดาบูซ (AdaBoost)

การสอนให้ระบบเรียนรู้และจดจำลักษณะของ Haar-like Feature เพื่อแยกแยะว่า “ใช่” หรือ “ไม่ใช่” ภาพใบหน้ามนุษย์นั้น Viola และ Jones ได้นำเอาวิธี AdaBoost หรือ Adaptive Boosting ซึ่งหลักการของ AdaBoost เป็นกระบวนการหาคุณลักษณะที่มีลักษณะใกล้เคียงกันและแตกต่างกันกับภาพที่นำเข้ามาสำหรับการจัดประเภทของภาพ โดยใช้การถ่วงน้ำหนักให้ส่วนต่างๆ ภายในภาพ ทั้งภาพตัวอย่างบวก (Positive) และภาพตัวอย่างลบ (Negative) เพื่อใช้หาลักษณะของวัตถุ (Object) ที่ “ใช่” และ “ไม่ใช่” ในลักษณะต่างๆ ได้ดังนี้

- 1) กำหนดค่าน้ำหนักให้กับคุณลักษณะที่วิ่งหาภายในภาพตัวอย่าง
- 2) หาบริเวณส่วนที่ต้องการ
- 3) เพิ่มค่าน้ำหนักให้กับส่วนที่เหลือเฉพาะส่วนที่ต้องการที่ยังไม่ได้แบ่งลักษณะไว้

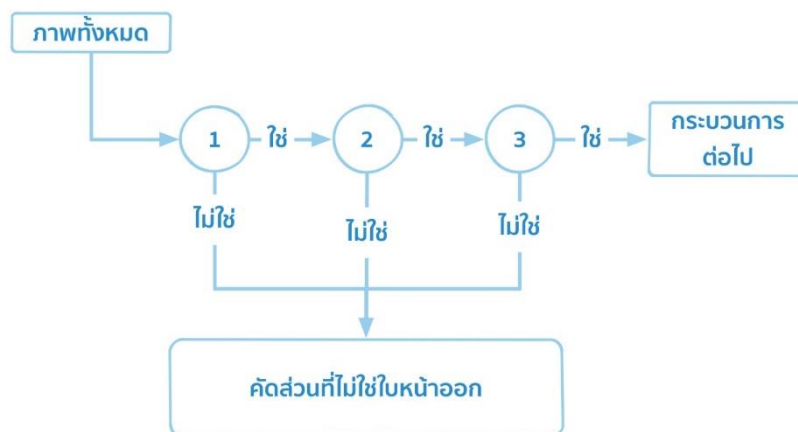
ทำวนเช่นนี้ไปเรื่อยๆ จนสุดท้ายนำบริเวณทั้งหมดมารวมกัน จะได้บริเวณของวัตถุ (Object) ที่ต้องการหาและลักษณะส่วนต่างๆ ภาพในภาพได้



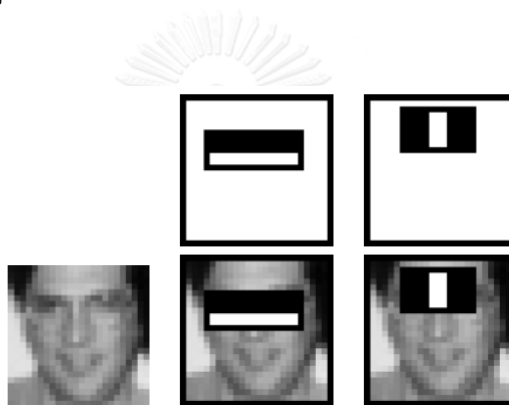
รูปที่ 2.10 การทำงานของเอดาบู้ซ [6]

### 2.1.6.3 ขั้นตอนการจำแนกรูปภาพด้วย Cascade classifier

ในขั้นตอนนี้เป็นกระบวนการตีความหมายของภาพหรือแยกแยะวัตถุที่ต้องการ โดยการแบ่งประเภทของภาพตามลักษณะภายในภาพ ซึ่งเริ่มต้นจากการตัดส่วนของ Sub-window ที่เป็น Negative หรือภาพที่ไม่มีวัตถุที่ต้องการค้นหาออกไปก่อนแล้วจากนั้นค่อยใช้ส่วนที่เป็น Positive วิ่งวนภายในภาพหากไม่เจอลักษณะที่ตรงกันก็จะเปลี่ยนลักษณะการตรวจจับภายใน Sub-window หากเจอลักษณะที่ตรงกัน ก็จะเปลี่ยนเป็นลักษณะที่มีความซับซ้อนมากขึ้น แต่วิธีการนี้จะทำให้มีโอกาสที่จะเกิดผลบวกหลง (False positive) น้อย ดังรูปที่ 2.11 แต่บางครั้งก็จะมีโอกาสเกิดผลลบหลง (False negative) ค่อนข้างมาก



รูปที่ 2.11 ขั้นตอนการจำแนกประเภทภาพใบหน้า

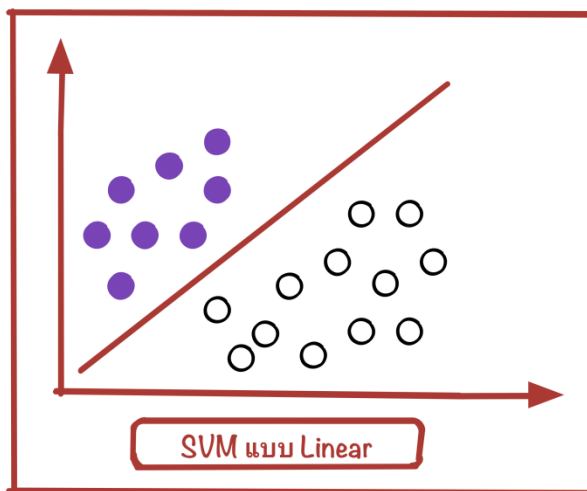


รูปที่ 2.12 ตัวอย่างการเรียนรู้ของเอดาบuchs [6]

จากรูปที่ 2.12 เป็นผลลัพธ์เบื้องต้นสำหรับการค้นหาภาพใบหน้า โดยการตีความหมายลักษณะบนใบหน้าที่มีความคล้ายคลึงกับสี่เหลี่ยม Haar-like Feature ซึ่งจะเห็นได้ว่าการเรียนรู้ของเอดาบuchs สองตัวอย่างแรกจะสามารถสังเกตได้ว่าในรูปแรกจะได้ลักษณะเด่น คือ บริเวณดวงตาคมมืดกว่าบริเวณอื่นๆ เช่น แก้ม เป็นต้น และในภาพที่สองจะสังเกตได้ว่าบริเวณสันจมูกจะสว่างกว่าบริเวณดวงตาทั้งสองข้าง

### 2.1.7 ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine)

ซัพพอร์ตเวกเตอร์แมชชีน [7] เป็นเทคนิคหนึ่งที่ได้รับคามนิยมอย่างแพร่หลายในการจำแนกประเภทรูปภาพ (Image Classification) โดยซัพพอร์ตเวกเตอร์แมชชีนมีหลักการ คือ ข้อมูลที่นำเข้ามาจะถูกแมปฟังก์ชัน (Mapping function) ไปยังอีกโดเมนที่เรียกว่าฟีเจอร์สเปซ (Feature space) และใช้เคอร์เนลฟังก์ชัน (Kernel Function) ในการวัดความคล้ายกันของข้อมูลในฟีเจอร์สเปซ จากนั้นจะสร้างไฮเปอร์เพลน (Hyper plane) มาแบ่งชุดข้อมูล 2 กลุ่มที่อยู่ในชุดฝึกสอน (Training dataset) โดยที่จะมีระยะห่างระหว่างขอบเขตทั้งสองกลุ่มมากที่สุด ดังรูปที่ 2.13



รูปที่ 2.13 ไฮเปอร์เพลนที่แบ่งชุดข้อมูลสองกลุ่ม

#### 2.1.8 ชุดไลบรารีของ OpenCV (Open Source Computer Vision library)

ชุดไลบรารีของ OpenCV [8] เป็นไลบรารีสำหรับใช้งานเรื่องการประมวลผลภาพและคอมพิวเตอร์วิทัศน์ และเป็นไลบรารีที่เป็นโอเพนซอร์ส ซึ่งถูกพัฒนาขึ้น โดยบริษัทอินเทลที่มีจุดเด่นในเรื่องการวิเคราะห์ภาพ ประมวลผลภาพทั้งที่เป็นภาพเคลื่อนไหวและภาพนิ่งหรือไฟล์วิดีโอ นอกจากนี้ OpenCV ยังเป็นชุดของคำสั่งไม่ได้เป็นตัวโปรแกรมสำเร็จรูป ดังนั้นการเรียนรู้ การใช้งานจึงสามารถเรียนรู้ได้รวดเร็ว ซึ่งเรียกส่วนที่ใช้เพียงชุดของคำสั่งที่ต้องการเท่านั้นและยังสามารถเขียนโดยใช้ได้หลากหลายภาษา เช่น ภาษา C, C++, C#, JAVA, Python เป็นต้น โดยที่ไลบรารีประกอบด้วย 4 ส่วนใหญ่ๆ ได้แก่

- CVCORE เป็นฟังก์ชันเบื้องต้นที่ใช้จัดการเกี่ยวกับ จุด ขนาด อาร์เรย์ หน่วยความจำ การประกาศตัวแปร เป็นต้น ตัวอย่างชุดคำสั่งที่นิยมใช้กัน ได้แก่ IplImage, CvMat
- CV ใช้ในการประมวลผลภาพและวิเคราะห์รูปภาพ โดยที่ฟังก์ชันส่วนใหญ่จะทำงานกับพิกเซลที่เป็นอาร์เรย์ 2 มิติ หรือที่เรียกกันว่าภาพ เช่น การหามุม การหาขอบ หรือการทำฮิสโตแกรม เป็นต้น
- HighGUI เป็นส่วนที่ใช้โหลดภาพ บันทึกภาพ ติดต่อกับอุปกรณ์ภายนอก กล้องวิดีโอ เม้าส์ แป้นพิมพ์ หรือการสร้างหน้าต่างเพื่อแสดงภาพและยกเลิกการใช้งานภาพ
- Machine Learning เป็นส่วนที่รวมฟังก์ชันทางสถิติ การแบ่งกลุ่มข้อมูล หรือการเรียนรู้ของเครื่องคอมพิวเตอร์

## 2.2 งานวิจัยที่เกี่ยวข้อง

ตารางที่ 2.1 แสดงจุดเด่น จุดด้อยของงานวิจัยที่เกี่ยวข้องพร้อมทั้งเทคนิควิธีที่ใช้

เรื่อง	เทคนิควิธี	จุดเด่น	จุดด้อย
Clothes Search in Consumer Photos via Color Matching and Attribute Learning [9]	<ul style="list-style-type: none"> <li>-เป็นการค้นหาความคล้ายของเสื้อผ้ากับภาพที่นำมาทดสอบ</li> <li>-เทคนิคที่ใช้ คือ สร้างคุณลักษณะจากสี โดยรวมแอททริบิวต์จำพวกชนิดเสื้อผ้า แขนสั้น แขนยาวเข้าไปสร้างเป็นคุณลักษณะด้วย</li> </ul>	<ul style="list-style-type: none"> <li>-ค้นหาเสื้อผ้า โดยเลือกจากสีของเสื้อผ้าเป็นส่วนสำคัญ</li> <li>-การนำแอททริบิวต์เข้ามาสร้างเป็นคุณลักษณะด้วยสามารถเพิ่มความถูกต้องในการค้นหาได้มาก</li> </ul>	<ul style="list-style-type: none"> <li>-ได้ระบุแอททริบิวต์ เช่น ชนิดของเสื้อผ้าก่อนการสร้างคุณลักษณะ</li> </ul>
A Sparse-Coding Based Approach to Clothing Image Retrieval [10]	<ul style="list-style-type: none"> <li>-ขั้นตอนการสร้างคุณลักษณะได้ใช้การรวมกันของเทคนิคสามอย่าง คือ ลักษณะเสื้อผ้า สี และเนื้อผ้า</li> <li>-ลักษณะของผ้าผู้เขียนได้แบ่งออกเป็น ความยาวของแขนเสื้อ กระดุม ปกเสื้อว่ามีหรือไม่มี ถ้ามียาวหรือสั้น</li> <li>-สีของเสื้อผ้าได้เปลี่ยนเป็นปริภูมิสีแบบ CIE-LAB แล้วสร้างฮิสโตแกรมในแต่ละช่องสี</li> <li>-เนื้อผ้าผู้เขียนได้ใช้ SIFT ในการค้นหาจุดสนใจร่วมกับ LBP และ Gabor</li> </ul>	<ul style="list-style-type: none"> <li>-การสร้างคุณลักษณะได้ใช้ลักษณะที่เห็นมาแบ่งแยกก่อนและใช้เทคนิคหลายๆ วิธีมารวมกัน ซึ่งทำให้มีรายละเอียดของการสร้างคุณลักษณะมากขึ้น</li> </ul>	<ul style="list-style-type: none"> <li>-ภาพเสื้อผ้าที่ใช้ในการทดสอบไม่ได้เป็นเสื้อผ้าที่มาจากตัวบุคคลสวมใส่ แต่มาจากแคตตาล็อกสินค้าสำเร็จรูป</li> <li>-ลักษณะของเสื้อผ้าได้ระบุไว้ก่อนว่าเป็นอะไร ส่วนสีและเนื้อผ้าไม่ได้ใช้จากเสื้อผ้าทั้งชิ้น แต่จะใช้การครอบภาพ ROI ในพื้นที่ตรงกลางของเสื้อผ้า</li> </ul>

ตารางที่ 2.1 แสดงจุดเด่น จุดด้อยของงานวิจัยที่เกี่ยวข้องพร้อมทั้งเทคนิควิธีที่ใช้ (ต่อ)

เรื่อง	เทคนิควิธี	จุดเด่น	จุดด้อย
Clothing Image Retrieval Using Color Moment [11]	<ul style="list-style-type: none"> <li>-ได้แปลงปริภูมิสีเป็นปริภูมิสีแบบเอชเอสวีแล้วใช้ฮิสโตแกรมของสีสร้างเป็นคุณลักษณะ</li> <li>-แบ่งตารางกริดขนาด 16x16 แล้วจึงใช้ฮิสโตแกรมไปทำในแต่ละกริด</li> <li>-ผลของค่าฮิสโตแกรมในแต่ละกริด คือ 32 ช่อง จะนำมาเรียงต่อกันเป็นหนึ่งคุณลักษณะ</li> </ul>	<ul style="list-style-type: none"> <li>-ใช้การคำนวณที่ไม่ยุ่งยากมากนักทั้งในเรื่องของการแบ่งตารางกริด และวิธีการสร้างคุณลักษณะจากเสื้อผ้า</li> </ul>	<ul style="list-style-type: none"> <li>-ใช้ทุกจุดพิกเซลในภาพในการนำมาสร้างเป็นฮิสโตแกรมของแต่ละตารางกริด ซึ่งสามารถต่อยอดการใช้ทุกจุดพิกเซลเป็นกลุ่มของจุดพิกเซลแล้วจึงหาค่าฮิสโตแกรม</li> </ul>
SNAPPER : Fashion Coordinate Image Retrieval System [12]	<ul style="list-style-type: none"> <li>-แบ่งเสื้อผ้าออกเป็น 4 ส่วน คือ เสื้อข้างนอก เสื้อข้างใน กางเกง และรองเท้า</li> <li>-สร้างคุณลักษณะจากสีร่วมกับจุดสนใจ</li> </ul>	<ul style="list-style-type: none"> <li>-แบ่งส่วนของเสื้อผ้าเพื่อสร้างคุณลักษณะ</li> </ul>	<ul style="list-style-type: none"> <li>-เสื้อผ้าที่นำมาค้นหาจะเป็นภาพที่ได้จากแคตตาล็อกสินค้าสำเร็จรูป ส่วนรูปภาพที่ผู้ค้นจะเป็นรูปภาพที่เป็นภาพบุคคล</li> <li>-แต่ละส่วนของเสื้อที่แบ่งออกไม่ได้นำมาสร้างเป็นคุณลักษณะเดียวกันแต่แบ่งออกเพื่อค้นหาในแต่ละส่วน</li> </ul>
Clothing Retrieval Based on Local Similarity with Multiple Images [13]	<ul style="list-style-type: none"> <li>-แบ่งโครงรูปภาพออกเป็น ส่วนแขนกับส่วนตัว</li> <li>-Local descriptor ได้ใช้วิธี RGB, LBP และ SIFT</li> </ul>	<ul style="list-style-type: none"> <li>-เลือกส่วนของรูปภาพจากหลายๆ ภาพแล้วนำมาสร้างเป็นคุณลักษณะ</li> <li>-เลือกส่วนของรูปภาพที่ไม่ต้องการนำมาสร้างเป็นคุณลักษณะได้</li> </ul>	

### บทที่ 3

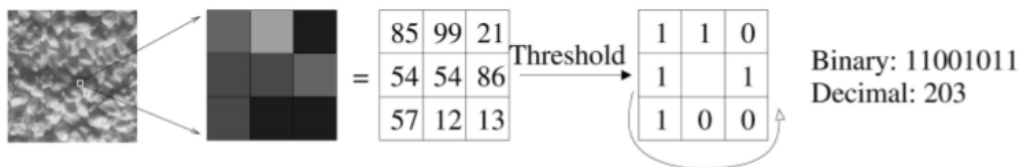
#### การดึงคุณลักษณะสำคัญของภาพ

ในส่วนนี้จะกล่าวถึงการดึงคุณลักษณะสำคัญท้องถิ่นของภาพ ซึ่งคุณลักษณะสำคัญท้องถิ่นเป็นวิธีการที่จัดอยู่ในกลุ่มของ Feature-based methods โดยสามารถคำนวณทางคณิตศาสตร์ได้จากการเปรียบเทียบพิกเซล (Pixel) แต่ละพิกเซลในภาพและพิกเซลเพื่อนบ้าน (Neighborhood pixel) เช่น ความเข้มของจุดภาพ (Intensity) และพื้นผิว (Texture) เป็นต้น ซึ่งคุณลักษณะสำคัญท้องถิ่นสามารถเป็นได้ทั้ง จุดพิกเซล ขอบภาพ หรือส่วนใดส่วนหนึ่งของภาพก็ได้ ซึ่งในที่นี้จะอธิบายคุณลักษณะสำคัญท้องถิ่นในการอธิบายลักษณะพื้นผิวของรูปภาพด้วยวิธีการ Local Binary Patterns (LBP) และ Local Directional Patterns (LDP) ส่วนกระบวนการดึงลักษณะเด่นด้วยการหาจุดสนใจบนภาพที่เรียกว่า Interest Points ด้วยวิธีการ Scale Invariant Feature Transform (SIFT) และ Speeded-Up Robust Features (SURF)

#### 3.1 การอธิบายลักษณะพื้นผิวของภาพ

##### 3.1.1 Local Binary Patterns (LBP)

Local Binary Patterns [14] เป็นหนึ่งในวิธีที่ดีที่สุดในการอธิบายและจัดการพื้นผิวของรูปภาพ ซึ่งได้ถูกนำไปใช้อย่างแพร่หลายเพราะง่ายต่อการคำนวณที่ไม่ยุ่งยากและใช้เวลาในการคำนวณน้อยอีกด้วย ซึ่งวิธีการคำนวณนั้นเป็นการทำในระดับพิกเซลของรูปภาพ ณ จุดพิกเซลใดๆ จะกำหนดเส้นขีดแบ่งขนาด  $3 \times 3$  พิกเซล โดยให้พิกเซลที่เป็นจุดศูนย์กลางเป็นค่าของเส้นขีดแบ่งผลที่ได้จะอยู่ในรูปของตัวเลขไบนารี คือ 0 กับ 1 ดังรูปที่ 3.1 แสดงวิธีการคำนวณหาค่าของไบนารีตามวิธีของ Local Binary Patterns แล้วแปลงกลับเป็นเลขฐานสิบ ผลที่ได้จากเลขฐานสิบจะเป็นค่าตัวเลขใหม่ให้กับจุดพิกเซลตรงกลางและทำทุกๆ จุดในภาพจะได้ดังรูปที่ 3.2



รูปที่ 3.1 วิธีการคำนวณ Local Binary Patterns [14]





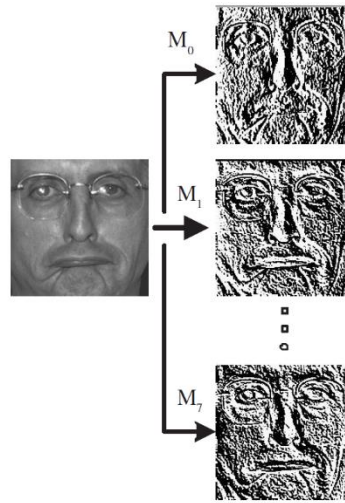
รูปที่ 3.2 ตัวอย่างรูปหลังจากการคำนวณด้วย Local Binary Patterns [14]

### 3.1.2 Local Directional Patterns (LDP)

Local Directional Patterns [15] เป็นคุณลักษณะสำคัญท้องถิ่นที่ถูกใช้มากในเรื่องของการวิเคราะห์และรู้จำภาพใบหน้า [16] ซึ่งวิธีการนี้จะเปรียบเทียบโครงร่างของภาพโดยพื้นฐานแล้วมีลักษณะคล้ายกันกับ Local Binary Pattern (LBP) แต่เปลี่ยนเป็นวิธีการแยกลักษณะเด่นเท่านั้น โดยมีวิธีการ คือ จะแบ่งโครงร่างใบหน้าออกเป็น 8 ทิศทาง ซึ่งแต่ละทิศทางจะผ่านตัวกรองต่างกัน ในรูปที่ 3.3 และผลลัพธ์ของภาพต้นแบบหลังผ่านตัวกรองทั้ง 8 ทิศทาง ในรูปที่ 3.4

$\begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix}$	$\begin{bmatrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{bmatrix}$	$\begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}$	$\begin{bmatrix} 5 & 5 & -3 \\ 5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}$
East $M_0$	North East $M_1$	North $M_2$	North West $M_3$
$\begin{bmatrix} 5 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & -3 & -3 \end{bmatrix}$	$\begin{bmatrix} -3 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix}$	$\begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix}$	$\begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{bmatrix}$
West $M_4$	South West $M_5$	South $M_6$	South East $M_7$

รูปที่ 3.3 Filter mask ทั้ง 8 ทิศทาง [15]

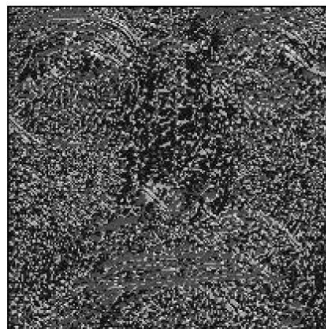


รูปที่ 3.4 ภาพหลังผ่าน Filter mask แต่ละทิศทาง [15]

ภาพที่ได้มาข้างต้นจะมีคุณลักษณะสำคัญแตกต่างกันออกไป ซึ่งในแต่ละคุณลักษณะสำคัญย่อยๆ ทั้ง 8 คุณลักษณะ คือ ตะวันออก ทิศตะวันออกเฉียงเหนือ ทิศเหนือ ทิศตะวันตกเฉียงเหนือ ทิศตะวันตก ทิศตะวันตกเฉียงใต้ ทิศใต้ และทิศตะวันออกเฉียงใต้ จะมีลำดับความสำคัญไม่เท่ากันทั้งหมด ดังนั้นจะต้องจัดอันดับความสำคัญของแต่ละคุณลักษณะสำคัญย่อยที่มีค่ามากที่สุด  $k$  ตัวให้มีค่าเท่ากับหนึ่งในการนำไปสร้างเป็นเลขฐานสองขนาด 8 บิต ส่วนคุณลักษณะสำคัญย่อยตัวอื่นๆ  $(8-k)$  จะกำหนดให้เป็นค่าศูนย์ จากรูปที่ 3.5 จะเห็นได้ว่าค่าของพิกเซล 50 จะถูกเปลี่ยนเป็นค่าของ LDP คือ 19 รูปที่ 3.6 แสดงภาพที่ผ่านการแปลงค่าทั้ง 8 ทิศทาง

85	32	26	Mask Index	$m_7$	$m_6$	$m_5$	$m_4$	$m_3$	$m_2$	$m_1$	$m_0$
53	50	10	Mask Value	161	97	161	537	313	97	-503	-393
60	38	45	Rank	6	7	5	1	4	8	2	3
			Code Bit	0	0	0	1	0	0	1	1
			LDP Code	19							

รูปที่ 3.5 ผลการจัดอันดับและแปลงเป็นเลขฐานสิบของจุดภาพ [15]



รูปที่ 3.6 ผลของภาพที่ได้จากการคำนวณ LBP [15]

## 3.2 การหาจุดสนใจของภาพ

กระบวนการหาจุดสนใจของภาพได้ถูกนำเสนอโดย Lowe [17] เป็นวิธีการที่เรียกว่า Scale Invariant Feature Transform หรือ SIFT ซึ่งหลังจากนั้น Bay ได้นำมาพัฒนาต่อด้วยการปรับแก้ไข SIFT ที่มีปัญหาในเรื่องของความเร็วในการทำงานที่ค่อนข้างช้า ให้มีการทำงานได้เร็วมากขึ้น และได้นำเสนอวิธีการที่เรียกว่า Speeded-Up Robust Features หรือ SURF โดยทั้งสองวิธีการมีขั้นตอนการทำงานหลักๆ อยู่ 2 ขั้นตอน คือ การหาจุดสนใจ (Keypoints) ของภาพและขั้นตอนการหาคุณสมบัติ (Descriptor) ของจุดสนใจจุดนั้นๆ

### 3.2.1 การหาจุดสนใจด้วย SIFT (Scale Invariant Feature Transform)

Scale Invariant Feature Transform [17] เป็นวิธีการจับคู่ภาพที่มีลักษณะคล้ายกันหรือเหมือนกันของภาพที่มีขนาดเท่ากัน ถ้าใช้วิธีการตรวจจับขอบหรือมุมของภาพก็ยังสามารถค้นหาความคล้ายได้ แต่ภาพในความจริงจะมีขนาดที่ไม่เท่ากันหรือภาพถูกหมุนเปลี่ยนทิศทางจะไม่สามารถค้นหาได้ ดังนั้น SIFT จึงได้ถูกพัฒนาขึ้นสำหรับการเปรียบเทียบความคล้ายกันของภาพ โดยที่ SIFT เป็น กระบวนการค้นหาและแยกคุณลักษณะเด่นที่คงที่ภายในภาพ ด้วยวิธีการหาจุดสนใจ (Interest points) ในภาพที่เรียกว่า Keypoints โดยไม่ขึ้นกับขนาดและทิศทางของวัตถุในภาพ ดังนั้นข้อดีของ SIFT ก็คือ ถ้าภาพถูกหมุน ขยาย หรือมูมมองและแสงเปลี่ยนไปเพียงเล็กน้อย วิธีการของ SIFT ก็ยังสามารถค้นหาจุดสนใจภายในภาพได้ ขั้นตอนการทำงานของ SIFT จะประกอบด้วยขั้นตอนสำคัญดังต่อไปนี้

#### 3.2.1.1 การเปลี่ยนแปลงระดับของภาพ

การหาลักษณะเด่นของภาพที่ไม่ได้ขึ้นกับขนาดและทิศทาง จะใช้วิธีการลดขนาด (Scale) และลดรายละเอียด (Blur) ของภาพให้มีขนาดเล็กลงไปเรื่อยๆ ซึ่งสามารถกำหนดได้ว่าภาพต้นแบบจะต้องลดขนาดและรายละเอียดลงเท่าไรก็ได้ แต่ตามที่ได้มีการทดลองวิธีการของ SIFT เพื่อกำหนดค่ามาตรฐานไว้ที่ลดขนาด 4 ระดับและลดรายละเอียด 5 ระดับ



รูปที่ 3.7 ภาพที่ได้จากการลดขนาด 4 ระดับและลดรายละเอียด 5 ระดับ [18]

ขั้นตอนการลดรายละเอียดภาพจะใช้วิธีการเบลอภาพ โดยเรียกการเบลอภาพนั้นจะอ้างถึงการทำคอนโวลูชันของฟังก์ชันเกาส์เซียน (Gaussian Function) ไปในแต่ละชั้น (Octave) ซึ่งในแต่ละชั้นก็จะมีรูปหลายๆ ระดับการเบลอ โดยมีขนาดรูปที่เท่ากันในแต่ละระดับการเบลอ จากนั้นจะค่อยเพิ่ม  $\sigma$  (Scale parameter) ทำให้ภาพเบลอมากขึ้น แล้วทำซ้ำกับขั้นต่อไปเรื่อยๆ โดยขั้นต่อไปก็จะมีขนาดของรูปเป็นครึ่งหนึ่งของ Octave ก่อนหน้าตามสมการที่ (6) และสมการที่ (7)

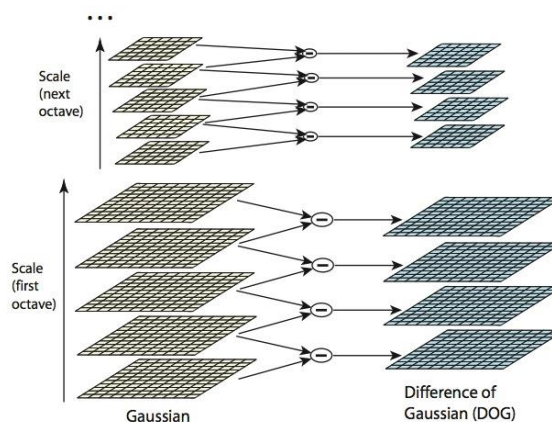
$$L(x, y, \sigma) = G(x, y, \sigma) \times I(x, y) \quad (6)$$

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (7)$$

- โดยที่ L คือ รูปภาพที่ได้จากการเบลอ  
 G ฟังก์ชันเกาส์เซียน คือ  
 I คือ ภาพตั้งต้น  
 x, y คือ พิกัดของจุดภาพ  
 $\sigma$  คือ พารามิเตอร์ที่ใช้ในการเบลอภาพ  
 $\times$  คือ การคอมโวลูชันระหว่างฟังก์ชันเกาส์เซียนกับรูปภาพตั้งต้นในทุกๆ จุดภาพ

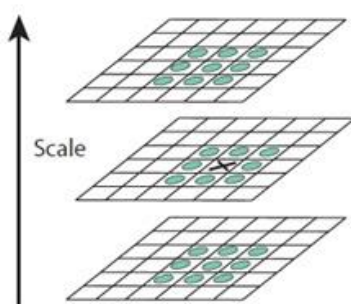
### 3.2.1.2 การกำหนดตำแหน่งจุดสนใจ

การใช้ Difference of Gaussian (DoG) เป็นวิธีที่นำมาระบุตำแหน่งของขอบหรือมุมในภาพ ซึ่งการหาขอบหรือมุมในภาพเป็นวิธีที่ดีในการนำมาใช้หาจุดสนใจในภาพ โดยวิธีการหาความต่างระหว่างภาพสองภาพที่ได้จากการลดขนาดและรายละเอียดในทุกๆ ระดับชั้นในรูปที่ 3.8



รูปที่ 3.8 การทำ Difference of Gaussian [18]

หลังจากทำ Difference of Gaussian แล้วในแต่ละระดับชั้นจะหาจุดสนใจที่มีค่ามากที่สุด (Maxima) หรือค่าน้อยที่สุด (Minima) โดยการเปรียบเทียบกับจุดรอบข้างทั้งแปดจุดและจุดที่อยู่ในระดับเดียวกันทั้งด้านข้างและด้านบน ซึ่งทำกับทุกๆ ระดับชั้น ดังรูปที่ 3.9 ที่จุด X จะเปรียบเทียบกับจุดรอบข้าง 8 จุด จากนั้นทำการเปรียบเทียบในระดับด้านบนและด้านล่าง ซึ่งรวมทั้งหมดในหนึ่งจุดที่ทำการเปรียบเทียบจะเท่ากับ  $8 + 9 + 9 = 26$  จุด ซึ่งถ้า X มีค่ามากที่สุดหรือน้อยที่สุดใน 26 จุด ก็จะทำให้ X เป็นจุดสนใจ



รูปที่ 3.9 การหาค่ามากที่สุดและน้อยสุดของจุด X กับจุดรอบข้างทั้ง 26 จุด [18]

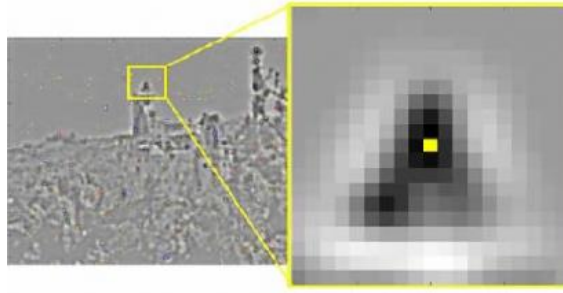
จุดสนใจในภาพที่ได้มาจากขั้นตอนที่ 2 จะพบว่ามีความคล้ายคลึงกันและยังเป็นจุดสนใจที่ไม่ดี เพราะบางจุดจะเป็นขอบภาพ บางจุดมีความต่างกันเพียงเล็กน้อย ดังนั้นจึงต้องกำจัดจุดสนใจที่ไม่ดีออก วิธีการกำจัดจุดสนใจที่ไม่ดี จะใช้การกระจายด้วยอนุกรมเทย์เลอร์ (Taylor series) ในการลดจำนวนจุดสนใจของภาพ ซึ่งเปรียบเทียบกับจุดสนใจที่มีค่ามากที่สุดหรือค่าน้อยสุดว่ามีความแตกต่างกันน้อยหรือไม่ ถ้าจุดสนใจใดๆ มีความแตกต่างกันน้อยก็กำจัดจุดสนใจนั้นออก หลังจากนั้นก็จะได้จุดสนใจที่เหมาะสมที่จะนำมาทำเป็นคุณลักษณะสำคัญ

### 3.2.2 การสร้างเอกลักษณ์ของแต่ละจุดด้วย SIFT (Interest Point Description)

ขั้นตอนต่อของ SIFT [18] เป็นการหาคุณสมบัติของจุดที่หาได้จากขั้นตอนก่อนหน้า โดยการหาคุณสมบัติของจุดใดๆ ที่ตรวจพบแล้วจะพิจารณาพื้นที่บริเวณรอบๆ จุดนั้น โดยมีขั้นตอนต่อไปนี้

#### 3.2.2.1 การกำหนดทิศทางของจุดสนใจ (Orientation Assignment)

ในขั้นตอนนี้จะได้ใส่ทิศทางให้กับทุกๆ จุดสนใจในภาพเพราะเมื่อภาพถูกรวม ย่อหรือขยายก็ยังสามารถจับคู่ภาพที่เหมือนกันได้อยู่ วิธีการ คือ คำนวณค่าทิศทางและค่าความเข้มทุกๆ จุดรอบๆ จุดสนใจนั้น แล้วนำไปคำนวณเพื่อสร้างทิศทาง ซึ่งทิศทางที่ได้จะมีหลากหลายทิศทางในที่นี้จะเลือกใช้ทิศทางที่มีค่าของผลรวมมากที่สุดนำมาแสดงรูปที่ 3.10 แสดงขนาดของจุดรอบๆ จุดสนใจนั้น



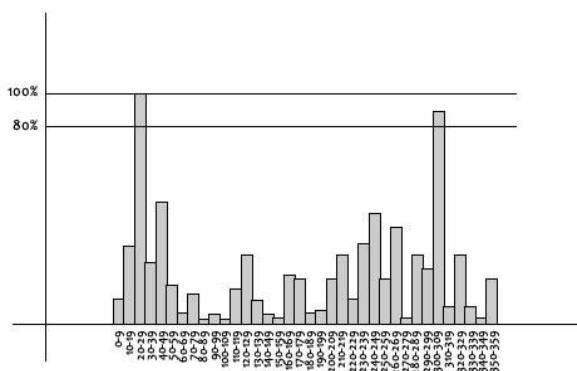
รูปที่ 3.10 ตัวอย่างจุดสนใจ [18]

สมการคำนวณทิศทางและคำนวณค่าความเข้มของจุดสีได้จากสมการที่ (8) และสมการที่ (9) โดยที่  $L$  คือ รูปภาพตั้งต้น ส่วน  $x, y$  คือ ตำแหน่งของจุดสนใจในภาพนั้นๆ

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (8)$$

$$\theta(x, y) = \tan^{-1} \left( \frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \right) \quad (9)$$

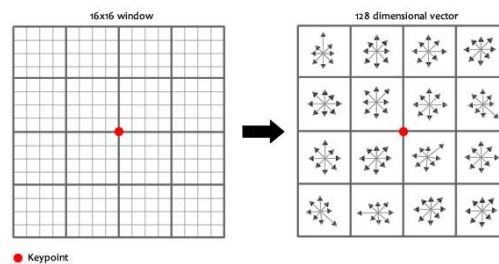
ในหนึ่งจุดสนใจจะพบว่าได้ทิศทางที่หลายทิศ ดังนั้นวิธีการที่จะระบุว่าจุดสนใจนั้นมีทิศทางอยู่ในทางไหนมากที่สุด จะใช้ฮิสโตแกรมในการรวบรวมความเข้มว่าในแต่ละทิศทางมีความเข้มเท่าใด ซึ่งในที่นี้ได้แบ่งช่วง (bin) ของฮิสโตแกรมไว้ที่ 36 ช่วง แต่ละช่วงมีองศาอยู่ที่ 10 องศา ดังรูปที่ 3.11 จะพบว่าแท่งที่สามมีองศาระหว่าง 20 - 29 องศา มีจำนวนความถี่ของความเข้มสูงที่สุด จึงถูกนำมาใช้เป็นทิศทางของจุดสนใจ แต่พบว่าองศาที่ 300 - 309 องศาที่สูงเกิน 80 เปอร์เซ็นต์ เหมือนกัน ดังนั้นจึงนำไปสร้างเป็นอีกหนึ่งจุดสนใจของภาพ



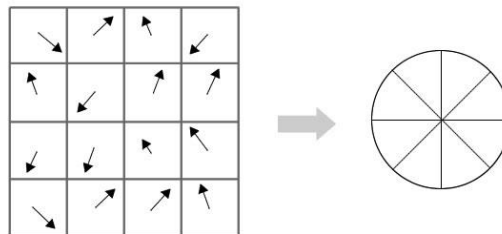
รูปที่ 3.11 ค่าฮิสโตแกรมทั้ง 36 ช่วง [18]

### 3.2.2 การอธิบายลักษณะเด่นของภาพ (Keypoint Descriptor)

จะกำหนดจุดรอบๆ จุดสนใจไว้ที่  $16 \times 16$  พิกเซล (พิกเซล คือ บริเวณที่จะทำการเก็บข้อมูลทิศทาง ซึ่งจะมีขนาดเท่ากับ  $1.5 \times \sigma$ ) และทำการแบ่งออกเป็น  $4 \times 4$  กลุ่ม ทั้งหมด 16 ชุด ซึ่งในแต่ละชุดจะคำนวณขนาดและทิศทางแล้วนำมาสร้างเป็นฮิสโตแกรมที่มีขนาด 8 bin (แกน x แบ่ง 8 ช่วง ช่วงละ 45 องศา) โดยขนาดของฮิสโตแกรมในแต่ละส่วนจะขึ้นกับ ขนาดคูณกับตัวถ่วงน้ำหนัก (ระยะที่ห่างจาก Keypoint)



รูปที่ 3.12 จุดรอบๆ จุดสนใจ [18]



รูปที่ 3.13 การเปลี่ยนค่าของทิศทางเป็นฮิสโตแกรม [18]

เมื่อทำการคำนวณหาค่าของฮิสโตแกรมทั้งหมด  $4 \times 4$  กลุ่ม ครบแล้วจะพบว่าขนาดของคุณลักษณะสำคัญจะมีขนาดเป็น  $4 \times 4 \times 8 = 128$  คุณลักษณะ ซึ่งจะนำไปใช้เป็นคุณลักษณะสำคัญ หรือ Feature vector ของแต่ละจุดสนใจต่อไป

### 3.2.3 การหาจุดสนใจด้วย SURF (Speeded-Up Robust Features)

Speeded-Up Robust Features [19] ได้รับการพัฒนามาจาก SIFT ในเรื่องของความเร็วในการคำนวณและความแม่นยำของการจับคู่ของภาพที่คล้ายกัน โดยหลักการของ SURF นั้นยังคล้ายวิธีของ SIFT วิธีนี้ได้ปรับปรุงการค้นหาจุดสนใจด้วยการใช้ Hessian matrix และเทคนิคการรวมภาพ (Integral image) ในการเพิ่มความเร็ว

### 3.2.3.1 เทคนิคการรวมภาพ (Integral Image)

SURF ได้นำเทคนิคการรวมภาพมาใช้เพื่อเพิ่มความเร็วในการคำนวณ ซึ่งวิธีการนี้สามารถคำนวณขนาดของพื้นที่สี่เหลี่ยมได้ทันทีหลังจากที่มีการใส่ข้อมูลที่ต้องการเข้ามา ทำให้ลดเวลาการคำนวณพื้นที่สี่เหลี่ยมของรูปภาพที่ต้องการได้ ตามสมการที่ (10)

$$II(x, y) = \sum_{x' \leq x, y' \leq y} I(x', y') \quad (10)$$

Original	Integral	Original	Integral
5 2 3 4 1	5 7 10 14 15	5 2 3 4 1	5 7 10 14 15
1 5 4 2 3	6 13 20 26 30	1 5 4 2 3	6 13 20 26 30
2 2 1 3 4	8 17 25 34 42	2 2 1 3 4	8 17 25 34 42
3 5 6 4 5	11 25 39 52 65	3 5 6 4 5	11 25 39 52 65
4 1 3 2 6	15 30 47 62 81	4 1 3 2 6	15 30 47 62 81

$5 + 2 + 3 + 1 + 5 + 4 = 20$ 
 $5 + 4 + 2 + 2 + 1 + 3 = 17$ 
 $34 - 14 - 8 + 5 = 17$

รูปที่ 3.14 ตัวอย่างการคำนวณพื้นที่ที่ต้องการ [20]

### 3.2.3.2 Hessian Detector

Hessian matrix ได้ถูกนำมาประยุกต์ใช้ในการหาจุดสนใจในภาพ โดยคิดจากค่าของ Determinant ซึ่งค่านี้จะนำมาใช้แยกค่าสูงสุดและค่าต่ำสุด

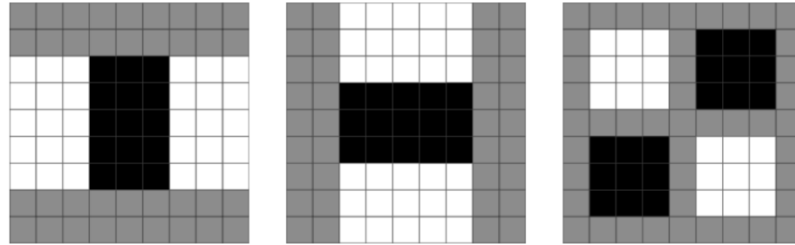
$$H(x, \sigma) \begin{bmatrix} \hat{L}_{xx}(x, \sigma) & \hat{L}_{xy}(x, \sigma) \\ \hat{L}_{xy}(x, \sigma) & \hat{L}_{yy}(x, \sigma) \end{bmatrix} \quad (11)$$

$$Det(H_{approx}) = \hat{L}_{xx}\hat{L}_{yy} - (0.9\hat{L}_{xy})^2$$

เมื่อ  $L_{xx}(x, \sigma)$  คือ การคำนวณคอนโวลูชันระหว่างอนุพันธ์ย่อยอันดับสองของเกาส์เซียนกับภาพตั้งต้นที่จุด  $x, y$  ใดๆ โดย SURF ใช้การประมาณค่าของ Hessian matrix ด้วยตัวกรองดังรูปที่ 3.15 คือ การประมาณค่าของอนุพันธ์ย่อยอันดับสองของเกาส์เซียน ซึ่งการประมาณค่านี้สามารถนำเทคนิคการรวมภาพมาใช้ เพื่อให้สามารถลดเวลาในการคำนวณลง โดยเวลาในการคำนวณนี้จะไม่ขึ้นกับขนาดของตัวกรอง ซึ่งค่าถ่วงน้ำหนักของ  $D_{xy}$  ในบริเวณพื้นที่สีดำจะมีค่าเป็น +1 และพื้นที่สีขาว



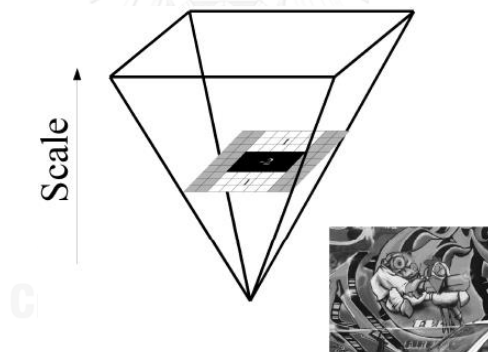
จะมีค่าเป็น -1 และค่าถ่วงน้ำหนักของ  $D_{xx}$  และ  $D_{yy}$  บริเวณพื้นที่สีดำจะมีค่าถ่วงน้ำหนักเป็น +2 ส่วนสีขาวจะมีค่าถ่วงน้ำหนักเป็น +1 ดังรูปที่ 3.15



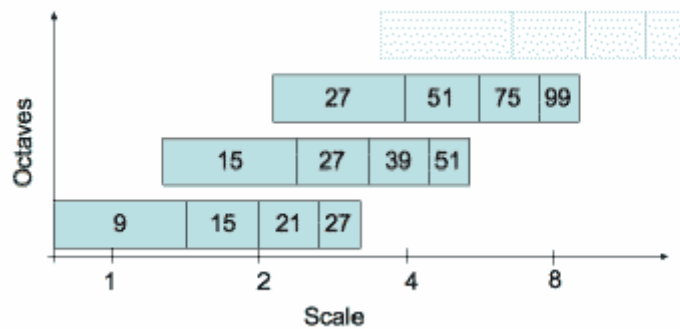
รูปที่ 3.15 ตัวอย่างตัวกรองในทิศทางของ x, y และ xy [21]

### 3.2.3.3 การเปลี่ยนแปลงระดับของภาพ

ขั้นตอนการเปลี่ยนแปลงระดับของภาพวิธีของ SURF จะใช้การเปลี่ยนขนาดของตัวกรองแทนขนาดของรูปภาพ จากรูปที่ 3.16 แสดงการเปลี่ยนขนาดของตัวกรองตามค่าต่างๆ ของอนุพันธ์ย่อยอันดับสองของเกาส์เซียนที่มีค่าสเกลต่างกันไป โดยได้มีการกำหนดค่าของแต่ละระดับชั้น (Octave) ไว้ที่ 5 ระดับชั้นและช่วงไว้ที่ 4 ช่วงในแต่ละระดับชั้น ดังรูปที่ 3.17



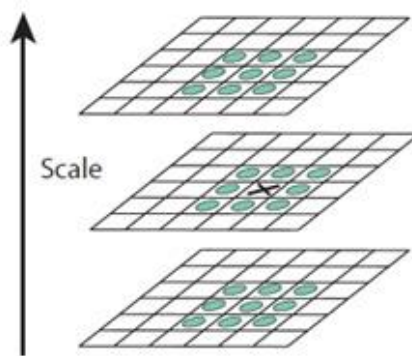
รูปที่ 3.16 การสเกลขนาดของตัวกรองที่มีขนาดต่างกัน [21]



รูปที่ 3.17 ตัวอย่างสเกลในแต่ละระดับชั้น [21]

### 3.2.3.4 การกำหนดตำแหน่งจุดสนใจ

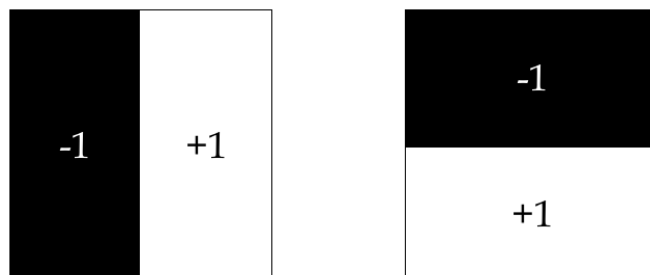
หลังจากได้ตรวจจับจุดสนใจในภาพแล้ว จะพบได้ว่าจุดสนใจที่ได้มามีปริมาณที่มาก ซึ่งอาจจะไม่ใช่จุดสนใจที่แท้จริงได้ ดังนั้นจะต้องทำการกำหนดเส้นขีดแบ่ง (Threshold) เพื่อใช้ในการกรองจุดสนใจออกไป โดยการเปรียบเทียบกับสเกลข้างเคียง ดังรูปที่ 3.18 จะเป็นการเปรียบเทียบกับพิกเซลทั้งหมด 26 พิกเซล แบ่งเป็นรอบข้างทั้ง 8 พิกเซล และ 9 พิกเซล ในระดับเดียวกันจากระดับเหนือกว่าและระดับต่ำกว่า โดยที่จุดพิกเซลที่พิจารณานั้นจะต้องเป็นพิกเซลค่ามากที่สุดหรือค่าน้อยสุด



รูปที่ 3.18 การเปรียบเทียบกับพิกเซลข้างเคียง [21]

### 3.2.4 การสร้างเอกลักษณ์ของแต่ละจุดด้วย SURF (Interest Point Description)

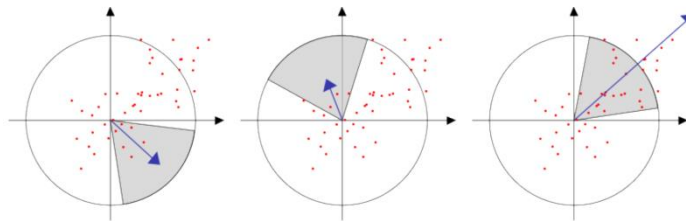
หลังจากที่ได้จุดสนใจในภาพแล้วขั้นตอนต่อไป คือ การสร้างเอกลักษณ์ของแต่ละจุดสนใจ ซึ่งในขั้นตอนนี้ได้นำ Haar Wavelet มาใช้ดังรูปที่ 3.19 โดย Haar Wavelet ที่ใช้ในการหาลักษณะเด่น จะมีสองขั้นตอนดังต่อไปนี้



รูปที่ 3.19 ตัวอย่าง Haar wavelet ในทิศทาง x และทิศทาง y

### 3.2.4.1 การกำหนดทิศทางของจุดสนใจ (Orientation Assignment)

วิธีการ คือ พิจารณาพิคเซลบริเวณรอบๆ จุดสนใจ โดย Haar wavelet ในทิศทาง  $x$  และ  $y$  ภายใต้วงกลมที่มีรัศมี  $6\sigma$  โดยค่าของ  $\sigma$  คือ ค่าสเกลของจุดสนใจที่กำลังพิจารณาอยู่ หลังจากที่ได้ค่ามาแล้ว จะนำค่าที่ได้มาทำการคำนวณค่าถ่วงน้ำหนักด้วยฟังก์ชันเกาส์เซียนขนาด  $2\sigma$  ซึ่งจะได้ค่าตอบสนองจุดสีแดง ดังรูปที่ 3.20 โดยหลังจากนั้นจะพิจารณาทิศทางของจุดที่ละ  $60^\circ$  จะได้ตามทิศทางของแต่ละพื้นที่ตามลูกศรสีน้ำเงิน โดยพื้นที่ใดมีขนาดของเวกเตอร์มากที่สุดจะให้เป็นตัวแทนทิศทางของจุดนั้น

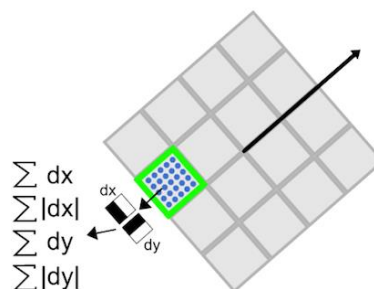


รูปที่ 3.20 ทิศทางของตัวกรอง Haar wavelet [21]

### 3.2.4.2 การอธิบายลักษณะเด่นของภาพ (Keypoint Descriptor)

การอธิบายลักษณะเด่นของภาพจะคำนวณมาจากจุดสนใจในภาพ โดยการสร้างสี่เหลี่ยมรอบจุดสนใจขนาด  $20\sigma$  โดยอ้างอิงจากสเกลที่กำลังพิจารณาอยู่ จากนั้นจะแบ่งพื้นที่ย่อยออกเป็น  $4 \times 4$  กลุ่ม ดังรูปที่ 3.21 หลังจากนั้นทำการคำนวณค่าของ Haar wavelet ในทิศทาง  $x$  และ  $y$  รูปที่ 3.19 ในแต่ละพื้นที่ย่อยดังแสดงในสมการที่ (12) แต่ละพื้นที่ย่อย 4 ช่องนี้เมื่อทำการหาคูณลักษณะของจุดสนใจแล้วก็จะมีค่า  $4 \times 4 \times 4 = 64$  ค่า ต่อหนึ่งจุดสนใจใดๆ ในภาพ

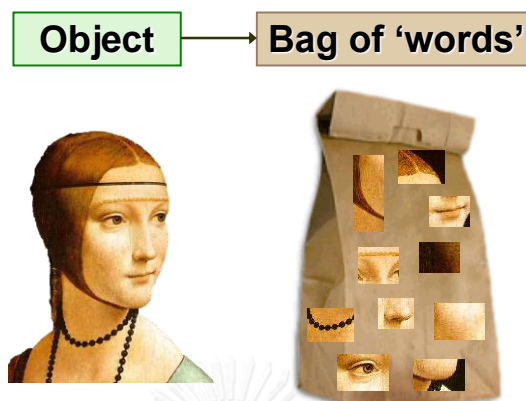
$$v_{subRegion} = (\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y|) \quad (12)$$



รูปที่ 3.21 ผลรวมของ Haar wavelet ของแต่ละพื้นที่ย่อย [21]

### 3.3 คุณลักษณะสำคัญครอบคลุม (Global Feature)

#### 3.3.1 Bag of Features (BoF)



รูปที่ 3.22 รูปภาพของคุณลักษณะสำคัญ [22]

รูปภาพของคุณลักษณะสำคัญ (Bag of Features) [23] เป็นวิธีที่ได้แนวคิดมาจากการจดจำและแบ่งประเภทข้อความที่เรียกว่า Bag of Words (BoW) โดยวิธีการ คือ จัดกลุ่มคำที่มีความหมายเหมือนกันอยู่ในกลุ่มเดียวกัน จากนั้นนำตัวแทนของคำในแต่ละกลุ่มมาสร้างเป็นพจนานุกรม เมื่อต้องการแบ่งประเภทข้อความใดๆ จะนำคำในข้อความนั้นมานับจำนวนความถี่ของคำที่ปรากฏอยู่ในพจนานุกรมว่าอยู่ในกลุ่มใดบ้าง แล้วจึงทำนายว่าข้อความนั้นเป็นประเภทอะไร ซึ่งแนวคิดนี้สามารถนำมาประยุกต์ใช้กับรูปภาพได้ ซึ่งเรียกว่า Bag of Visual Words หรือ Bag of Features และพจนานุกรมของภาพจะเรียกว่า Codebook ในแต่ละขั้นตอนการทำงานของ Bag of Features จะมีวิธีการดังต่อไปนี้

##### 3.3.1.1 การดึงคุณลักษณะสำคัญของจุดสนใจ

หลังจากขั้นตอนการค้นหาจุดสนใจด้วยวิธี Scale Invariant Feature Transform (SIFT) หรือวิธี Speeded-Up Robust Features (SURF) จะได้ตำแหน่งของจุดสนใจ (Interest Points) ของภาพออกมา ซึ่งขนาดของเวกเตอร์ที่ใช้ในการสร้างพจนานุกรมของภาพจะขึ้นอยู่กับขนาดของ KeyPoint Descriptor ถ้าใช้วิธี SIFT จะได้ขนาดเวกเตอร์เท่ากับ 128 หากใช้วิธีของ SURF จะได้ขนาดเวกเตอร์เท่ากับ 64 ต่อหนึ่งจุดสนใจในภาพ

##### 3.3.1.2 สร้างพจนานุกรมของจุดสนใจ

กระบวนการสร้างพจนานุกรม (Codebook) จะเป็นการจัดกลุ่ม (Clustering) ของจุดสนใจทั้งหมดในตัวอย่างข้อมูลฝึกสอน โดยมีหลักการ คือ จัดกลุ่มสนใจที่มีลักษณะคล้ายกันเอาไว้กลุ่ม

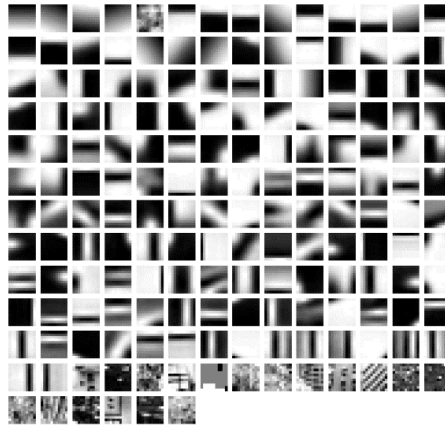
เดียวกัน วิธีการจัดกลุ่มสนใจที่มีลักษณะคล้ายกันนั้นจะนำ K-means เข้ามาจัดกลุ่ม โดยการวัดระยะทางแบบยุคลิดระหว่างเวกเตอร์ของจุดสนใจทั้งหมดในตัวอย่างข้อมูลฝึกสอน ดังสมการที่ (13) ซึ่งการกำหนดขนาดของ K จะต้องไม่มากเกินไปจะทำให้เกิด Over-fitting หรือน้อยเกินไปอาจจะทำให้เกิด Under-fitting ขึ้นได้ โดยในวิทยานิพนธ์นี้ ผู้วิจัยได้กำหนดขนาดของ K เท่ากับ 1,024 กลุ่ม

$$D(X, M) = \sum_{i=1}^k \sum_{j=1}^k (x_j - m_k)^2 \quad (13)$$

โดยที่ D คือ ค่าระยะทางแบบยุคลิดระหว่างจุดสนใจกับจุดกลางของกลุ่ม  
 k คือ จำนวนกลุ่มคลัสเตอร์  
 $x_j$  คือ จุดสนใจ  
 $m_k$  คือ จุดที่แบ่งกลุ่มคลัสเตอร์

จากสมการที่ (13) คือ สมการวัดระยะทางแบบยุคลิดที่ใช้ในการจัดกลุ่มใน K-means ซึ่งวิธีการจัดกลุ่มโดยใช้ K-means นั้นมีหลักการดังนี้

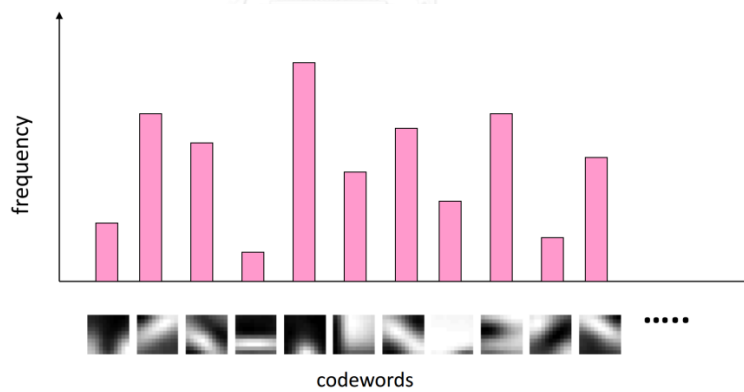
- 1) กำหนดค่าของ K หรือจำนวนกลุ่มที่ต้องการจะจัดกลุ่ม ในที่นี้ผู้วิจัยได้กำหนดจำนวนกลุ่มไว้ที่ 1,024 กลุ่ม
- 2) ทำการสุ่มค่าเวกเตอร์ตัวแทนกลุ่มเริ่มต้นจำนวน 1,024 ตัว เพื่อให้เป็นจุดศูนย์กลาง (Centroid) ของแต่ละกลุ่ม
- 3) เวกเตอร์ของจุดสนใจทั้งหมดจะถูกวัดระยะห่างระหว่างเวกเตอร์ตัวแทนกลุ่มของแต่ละกลุ่มกับเวกเตอร์ของจุดสนใจ เวกเตอร์ของจุดสนใจจะถูกให้อยู่คลัสเตอร์เดียวกันก็ต่อเมื่อเวกเตอร์ของจุดสนใจนั้นมีระยะห่างที่ใกล้มากที่สุด
- 4) คำนวณหาค่าเฉลี่ยใหม่ของเวกเตอร์ตัวแทนและวัดระยะทางของเวกเตอร์ของจุดสนใจแล้วจัดกลุ่มเวกเตอร์ของจุดสนใจใหม่
- 5) ทำจนกระทั่งเวกเตอร์ของจุดสนใจไม่มีการเปลี่ยนกลุ่มคลัสเตอร์



รูปที่ 3.23 ตัวอย่างพจนานุกรมของภาพ [22]

### 3.3.1.3 Histogram of Features

การสร้างคุณลักษณะสำคัญเพื่อใช้ในการสอนให้ระบบเรียนรู้สำหรับการจำแนกประเภทนั้น จะใช้วิธีดูค่าความถี่ของจุดสนใจในรูปภาพตั้งต้นเมื่อนำมาเปรียบเทียบกับพจนานุกรมของภาพ ดังนั้นขนาดของเวกเตอร์ที่ใช้ในการสอนให้ระบบเรียนรู้จะมีขนาดเท่ากับ 1,024 ดังรูปที่ 3.24 และต้องทำการ Normalize ค่าของฮิสโตแกรมให้อยู่ในช่วงของ 0 ถึง 1 ก่อนการสอนให้ระบบเรียนรู้

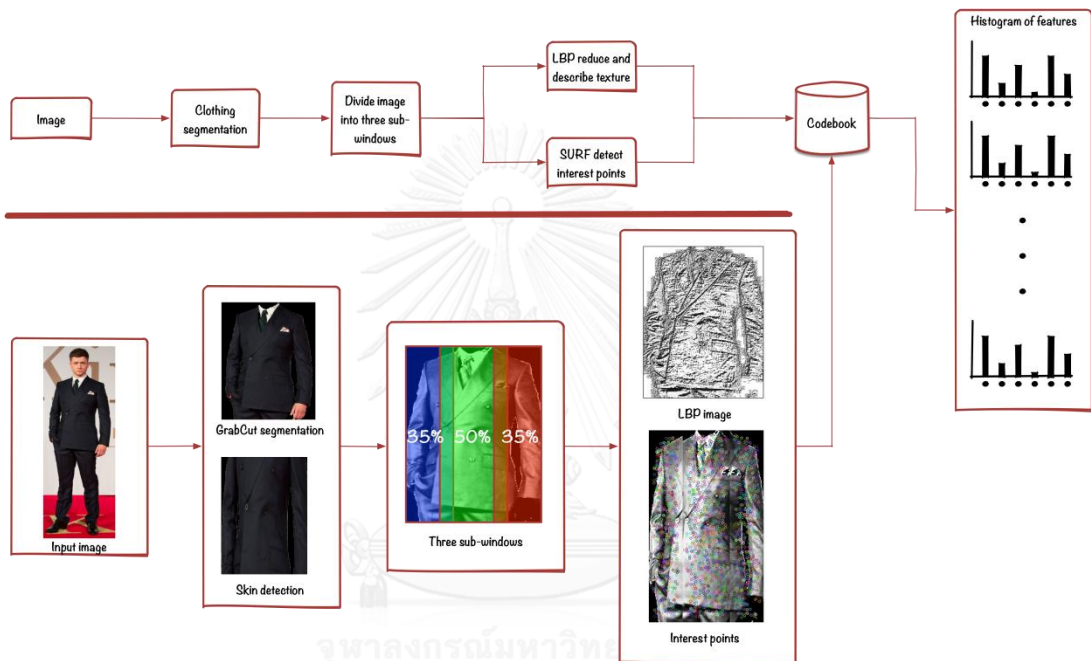


รูปที่ 3.24 ตัวอย่าง Histogram of Features [22]

## บทที่ 4

### ขั้นตอนการดำเนินงานและการออกแบบ

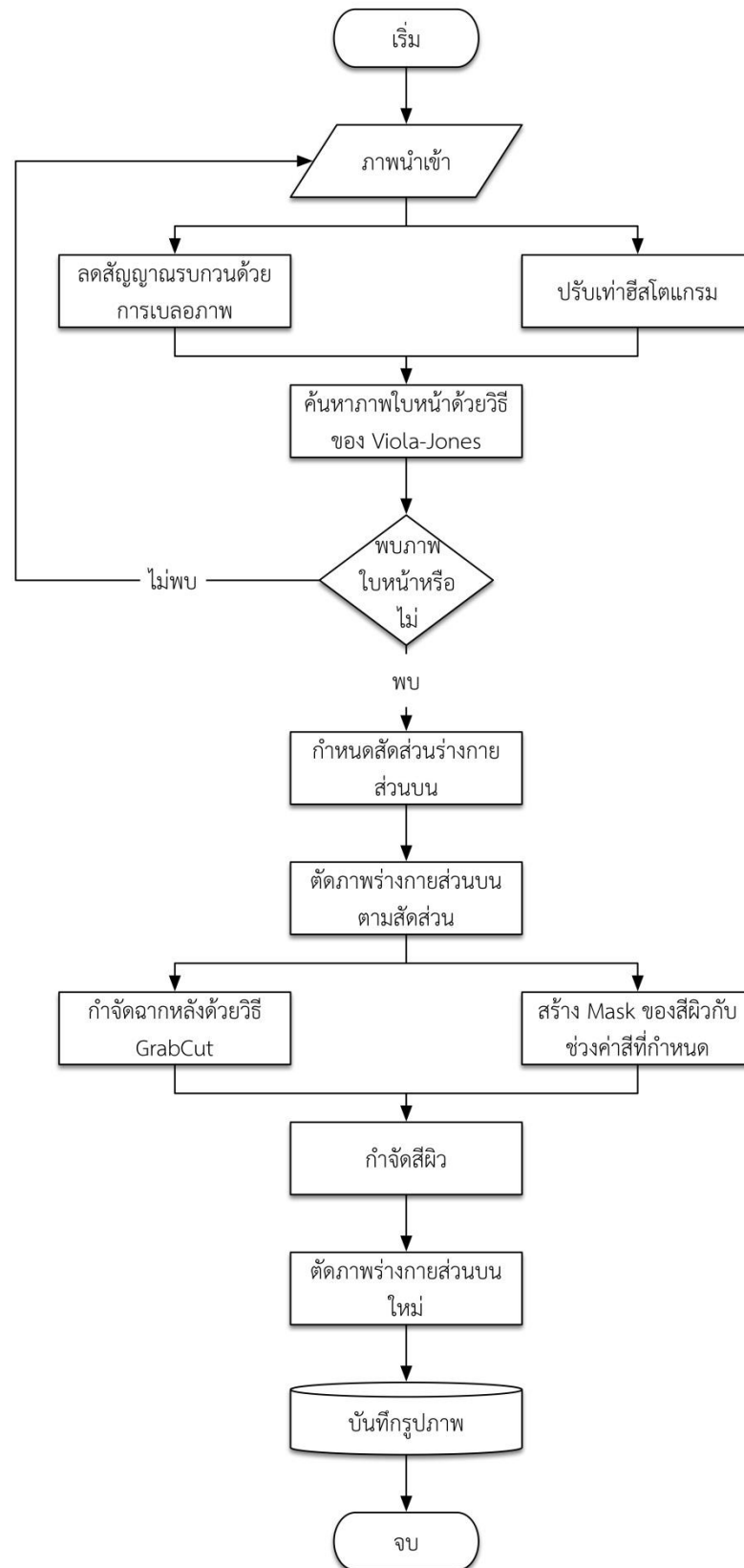
งานวิจัยนี้ได้นำเสนอวิธีการรู้จำชนิดของเสื้อผ้า ด้วยการจำแนกประเภท โดยได้แบ่งภาพรวมของวิธีการไว้ 3 วิธี คือ วิธีการแบ่งส่วนของเสื้อผ้า วิธีการสร้างคุณลักษณะสำคัญของเสื้อผ้าเพื่อสร้างโมเดลของเสื้อผ้า วิธีการสอนให้เครื่องเรียนรู้และการทดสอบ ดังรูปที่ 4.1 ภาพรวมของวิธีการรู้จำชนิดของเสื้อผ้า



รูปที่ 4.1 ภาพรวมของวิธีการรู้จำชนิดของเสื้อผ้า

#### 4.1 การแบ่งส่วนของเสื้อผ้า (Clothing segmentation)

จากรูปที่ 4.2 แสดงขั้นตอนการทำงานของแต่ละขั้นตอนในการแบ่งส่วนของเสื้อผ้า ซึ่งภาพที่นำเข้าจะผ่านการประมวลผลภาพเบื้องต้น (Image Preprocessing) ก่อน เช่น การกำจัดสัญญาณรบกวนด้วยวิธีเบลอภาพ ปรับขนาดของภาพ หรือการปรับปรุงแสงโดยใช้ฮิสโตแกรม เป็นต้น หลังจากนั้นเมื่อผ่านขั้นตอนวิธีการต่างๆ ในการแบ่งส่วนของภาพแล้ว รูปภาพของเสื้อก็จะถูกบันทึกลงในฐานข้อมูลเพื่อใช้ในขั้นตอนถัดไป



รูปที่ 4.2 แสดงแต่ละขั้นตอนของการแบ่งส่วนของเสื้อผ้า

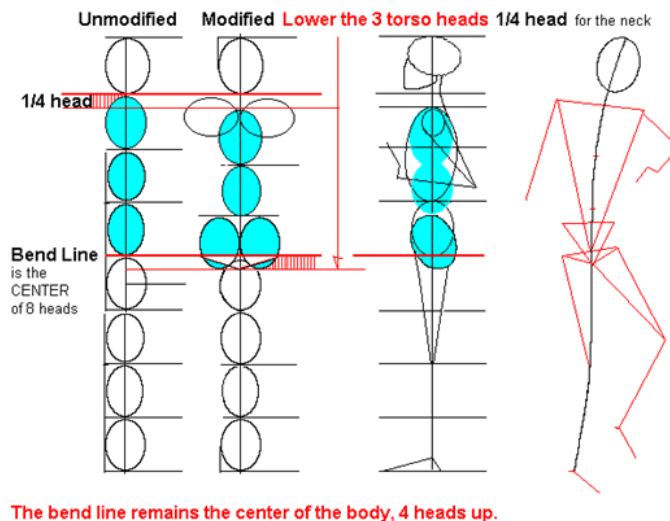


#### 4.1.1 การค้นหาใบหน้าและร่างกายส่วนบน

ในขั้นตอนแรกของการค้นหาส่วนที่เป็นเสื้อผ้าในภาพ วิธีที่ง่ายที่สุดและใช้กันอย่างแพร่หลายคือ การค้นหาภาพใบหน้าก่อนและสมมติว่าส่วนที่เป็นเสื้อผ้าจะอยู่ใต้ใบหน้าเสมอ โดยวิธีการค้นหาภาพใบหน้าในวิทยานิพนธ์นี้ได้ใช้วิธีของ Viola-Jones [6] ซึ่งได้อธิบายขั้นตอนการทำงานไว้ในบทที่สอง เมื่อได้ตำแหน่งและขนาดของใบหน้าในภาพจะต้องคำนวณขนาดของใบหน้าใหม่เพื่อนำไปเปรียบเทียบกับขนาดสัดส่วนของหัว จึงสามารถกำหนดขนาดของร่างกายส่วนบนได้ โดยเปรียบเทียบกับสัดส่วนของหัวตามสมการที่ (14) จะแสดงวิธีการคำนวณสัดส่วนของหัวจากขนาดของใบหน้า เมื่อ  $H(x, y)$  คือ ขนาดของหัว และ  $f(x, y)$  ขนาดของใบหน้า

$$H(x, y) = \frac{1}{3} \left( \frac{f(x, y)}{2} \right) + f(x, y) \quad (14)$$

เมื่อได้ขนาดของหัวแล้ว วิธีที่ใช้สำหรับเปรียบเทียบกับขนาดของร่างกาย ผู้วิจัยได้ใช้สัดส่วนการเขียนแบบร่างของมนุษย์เป็นพื้นฐานในการกำหนดขนาดของร่างกายส่วนบน ซึ่งในรูปที่ 3.25 จะแสดงสัดส่วนแบบร่างในการเขียนแบบของมนุษย์



รูปที่ 3.25 สัดส่วนของการเขียนแบบร่างมนุษย์ [24]

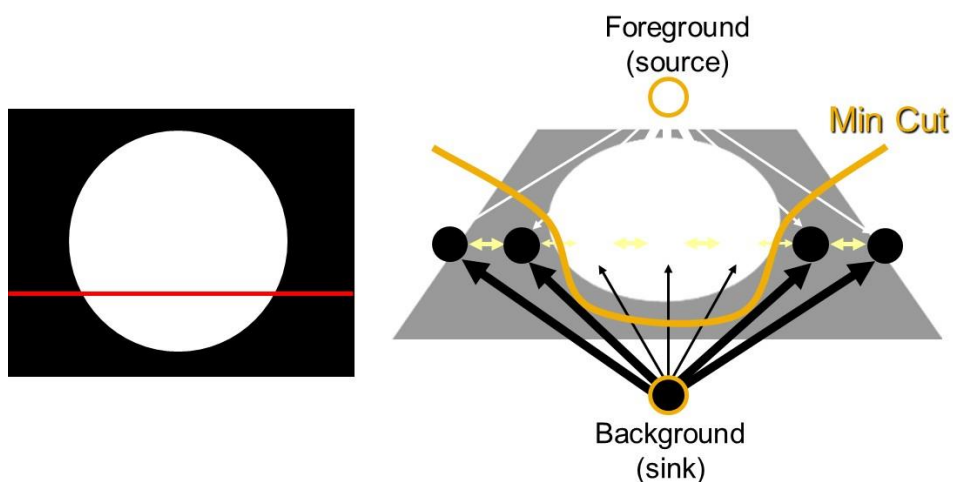
เมื่อได้ขนาดของร่างกายแล้วจะตรวจสอบว่าพบเสื้อผ้าตรงตำแหน่งของร่างกายส่วนบนหรือไม่ ถ้าพบจะทำการระบุตำแหน่งและขนาดลงในรูปภาพ ดังรูปที่ 4.3 แสดงการกำหนดสัดส่วนของร่างกายส่วนบน



รูปที่ 4.3 แสดงการกำหนดสัดส่วนของร่างกายส่วนบน [25]

#### 4.1.2 การแบ่งส่วนภาพฉากหน้าและฉากหลัง

เมื่อได้ตำแหน่งของเสื้อในภาพแล้วจะพบว่าภายในภาพมีส่วนที่ประกอบไปด้วย เสื้อผ้า สีผิว และฉากหลัง ถ้านำไปสร้างคุณลักษณะผลที่ได้อาจจะได้ค่าความถูกต้องที่ผิดพลาด ดังนั้นก่อนนำภาพไปสร้างคุณลักษณะจำเป็นต้องกำจัดสิ่งที่ไม่เกี่ยวข้องออกไปก่อน โดยในขั้นตอนนี้ได้นำเสนอการกำจัดฉากหลังด้วยวิธี GrabCut [26] มีวิธีการทำ คือ สร้างกรอบสี่เหลี่ยมรอบเสื้อและกำหนดว่าภายในกรอบสี่เหลี่ยมนี้เป็นฉากหน้าส่วนที่อยู่นอกกรอบให้เป็นฉากหลัง GrabCut จะมองพิกเซลแต่ละพิกเซลว่าเป็นโนดหนึ่งๆ และมีเส้นเชื่อมต่อกัน โดยที่ค่าแรงยึดติดของเส้นเชื่อมจะคำนวณมาจากการจัดกลุ่มของสี (Color clustering) ด้วยวิธี Gaussian Mixture Models (GMMs) และวัดค่าแรงยึดติดของเส้นเชื่อมในแต่ละพิกเซล ค่าแรงยึดติดที่มีค่าน้อยๆ GrabCut จะตัดเส้นเชื่อมนั้นออกและจัดกลุ่มของสีอีกครั้ง จนกระทั่งเส้นที่ตัดออกจะมาบรรจบกัน ดังรูปที่ 4.4 ซึ่งในโนดใดๆ จะต้องมีเส้นเชื่อมอย่างน้อยหนึ่งเส้น

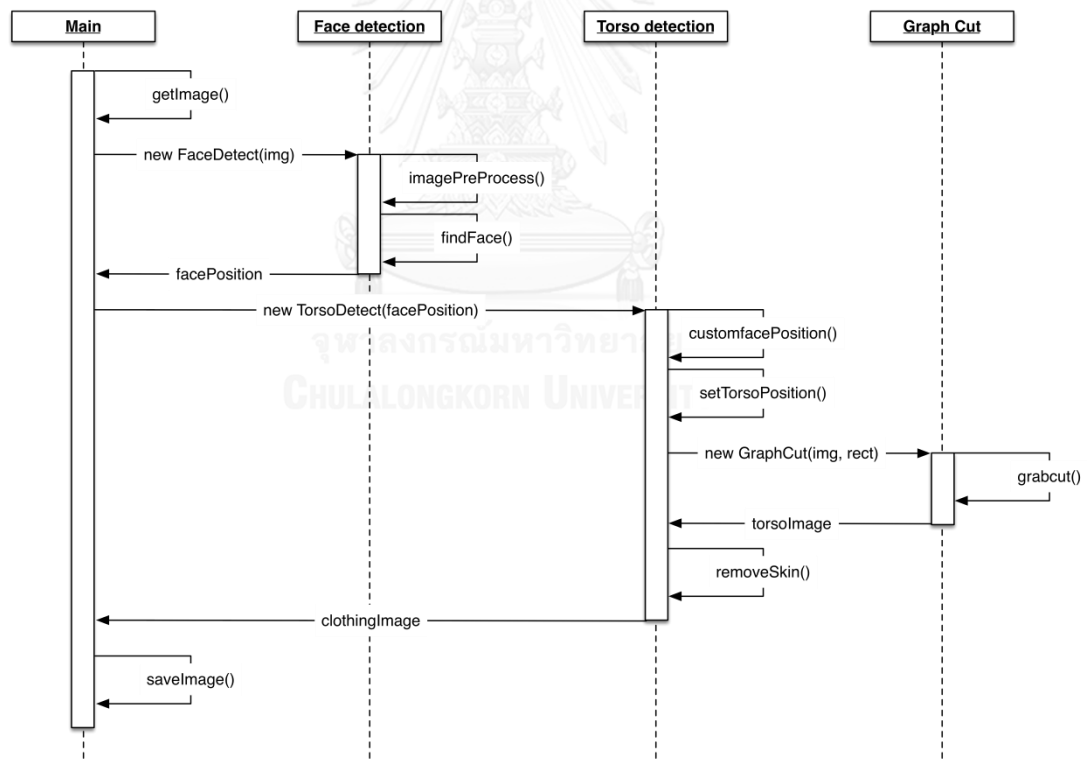


รูปที่ 4.4 ขั้นตอนการแยกฉากหน้าและฉากหลังด้วยวิธี GrabCut [27]



รูปที่ 4.5 ผลที่ได้จากการกำจัดฉากหลัง

จากรูปที่ 4.5 แสดงผลการใช้งาน GrabCut ในการกำจัดฉากหลังด้วยการระบุกรอบสี่เหลี่ยมที่เป็นฉากหน้าทำให้สามารถได้ภาพที่เป็นฉากหน้าได้อัตโนมัติ โดยไม่ต้องใช้มนุษย์ระบุพื้นที่ที่เป็นฉากหน้าและฉากหลังเองทำให้ขั้นตอนวิธีของ GrabCut เป็นที่นิยมอย่างมากในรูปภาพที่เป็นภาพนิ่ง



รูปที่ 4.6 แผนภาพลำดับของการแบ่งส่วนฉากหน้า

จากรูปที่ 4.6 ได้นำเสนอเป็นแผนภาพลำดับของกระบวนการค้นหาเสื้อผ้าและแบ่งส่วนฉากหน้าออกจากฉากหลัง ในตั้งแต่ขั้นตอนการค้นหาใบหน้า การกำหนดสัดส่วนร่างกายส่วนบนจนกระทั่งการแบ่งแยกโดยใช้ GrabCut ผลที่ได้จะบันทึกลงในฐานข้อมูล

#### 4.1.3 การแบ่งส่วนสีผิว

หลังจากแบ่งแยกฉากหน้าและฉากหลังจากการใช้ GrabCut ร่วมกับกรอบสี่เหลี่ยมที่คาดว่าจะเป็นเสื้อผ้าภายในรูปภาพแล้ว จะพบว่าภาพที่ได้มาเป็นฉากหน้าอาจจะไม่ได้มีเพียงแค่เสื้อผ้าเพียงอย่างเดียวแต่จะรวมถึงสีผิวของมนุษย์ด้วย ถ้านำไปสร้างเป็นคุณลักษณะจะทำให้ผลที่ได้จากการรู้จำผิดพลาด ดังนั้นในส่วนนี้จะกล่าวถึงการกำจัดสีผิวของมนุษย์ออกจากรูปภาพก่อนนำไปสร้างเป็นคุณลักษณะด้วยวิธีการแบ่งแยกสี (Color segmentation) และดึงส่วนที่เป็นสีผิวออก ซึ่งปัญหาของการแบ่งแยกสีผิว คือ ถ้าช่วงของแสงมีการเปลี่ยนแปลงอาจจะทำให้ไม่สามารถกำจัดสีผิวออกได้ หรือในต่างพื้นที่ สิ่งแวดล้อมก็จะส่งผลต่อการกำจัดสีผิวเช่นกัน

Albiol et al. [28] ได้นำเสนอวิธีการกำหนดค่าสีของผิวมนุษย์ว่าอยู่ในช่วงที่เท่าไรที่ใช้ในการแบ่งแยกสีผิว ในงานวิจัยนี้ผู้วิจัยได้เปลี่ยนปริภูมิสีจากปริภูมิสีแบบอาร์จีบีเป็นปริภูมิสีแบบเอชเอสวี ก่อนการทดลองเพื่อง่ายต่อการกำหนดค่าของสีผิวมนุษย์ ผลที่ได้จากการทดลองพบว่าค่าสีที่เหมาะสมกับการกำจัดสีผิวมนุษย์ในงานวิจัยนี้จะอยู่ในช่วง  $H = [0, 42]$ ,  $S = [32, 235]$  และ  $V = [60, 255]$  และใช้สมการที่ (15) ในการสร้างรูปภาพเป็นภาพแบบไบนารีของสีผิวมนุษย์ โดยวิธีการแปลง คือ ถ้าพิกเซลใดๆ เป็นสีผิวจะกำหนดให้เป็นหนึ่ง แต่ถ้าไม่ใช่จะเป็นศูนย์

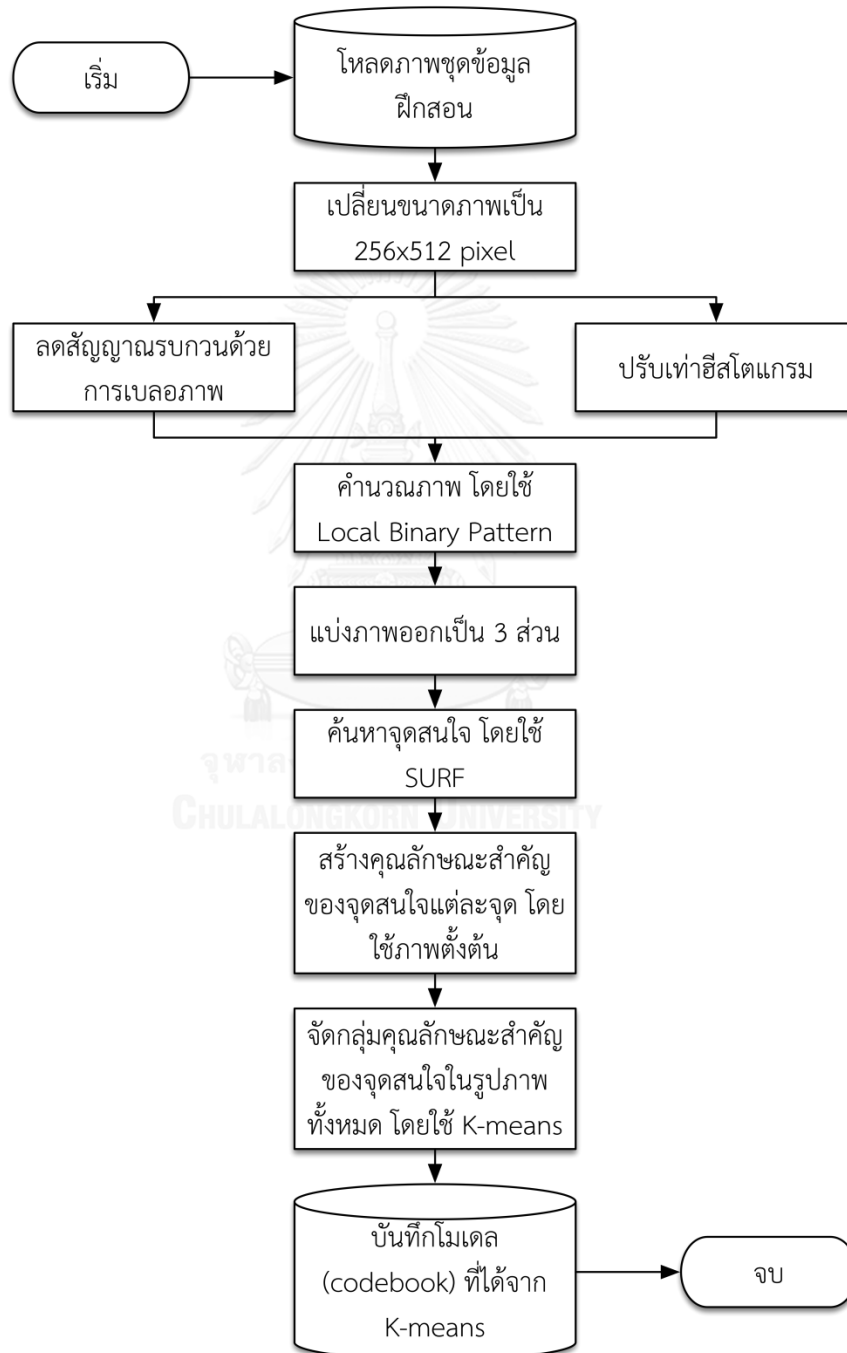
$$skin(x, y) = \begin{cases} 1, & \text{if a pixel is not in ranges} \\ 0, & \text{otherwise} \end{cases} \quad (15)$$



รูปที่ 4.7 ผลของการกำจัดสีผิวมนุษย์ในภาพเสื้อผ้า

#### 4.2 การดึงคุณลักษณะสำคัญของเสื้อผ้า (Features extraction)

จากรูปที่ 4.8 แสดงขั้นตอนในการดึงคุณลักษณะสำคัญของจุดสนใจในเสื้อผ้า เพื่อนำไปสร้างเป็นพจนานุกรมของภาพ ผลลัพธ์ที่ได้จากขั้นตอนนี้ คือ โมเดลของ K-means จะถูกบันทึกเพื่อนำไปใช้ในการสร้างคุณลักษณะสำคัญในการสอนระบบให้เรียนรู้



รูปที่ 4.8 การดึงคุณลักษณะสำคัญของเสื้อผ้าเพื่อนำไปสร้างเป็นพจนานุกรมของภาพ

#### 4.2.1 การจัดการข้อมูลนำเข้าก่อนการวิเคราะห์ (Data preprocessing)

ขั้นตอนการดึงคุณลักษณะสำคัญของเสื้อผ้าจำเป็นต้องมีการจัดการข้อมูลก่อนการวิเคราะห์ ทั้งขนาดของรูปภาพที่ได้มาและการกำจัดสัญญาณรบกวนต่างๆ ซึ่ง ในที่นี้ผู้วิจัยได้กำหนดขนาดเป็นของรูปภาพทั้งหมดในฐานข้อมูลเป็น 256x512 พิกเซล และได้ใช้การปรับเท่าฮิสโตแกรมเพื่อกระจายความเข้มของแสงตลอดทั้งภาพ

ในที่นี้ผู้วิจัยได้นำเสนอการแบ่งส่วนของรูปภาพที่จะนำไปดึงคุณลักษณะสำคัญออกเป็น 3 ส่วน คือ ด้านซ้ายของเสื้อผ้า ด้านขวาของเสื้อผ้าและตรงกลาง เหตุผลที่แบ่งออกเป็น 3 ส่วน คือ ผู้วิจัยคิดว่าพื้นที่ของเสื้อผ้ามีความสำคัญไม่เท่ากันในแต่ละด้านของเสื้อผ้า เช่น ในด้านตรงกลางเสื้อผ้าชนิดที่เป็นเสื้อโปโลจะแตกต่างจากเสื้อยืด เพราะมีกระดุม เสื้อเชิ้ตกระดุมยาวลงมาตลอด หรือเสื้อสูทที่มีเนคไทผูกอยู่ เป็นต้น ดังนั้นถ้าแบ่งส่วนออกไปและให้ความสำคัญกับส่วนที่มีความสำคัญมากเมื่อนำไปสร้างเป็นคุณลักษณะสำคัญน่าจะได้คุณลักษณะสำคัญที่ดีเช่นกัน

การแบ่งส่วนตามความสำคัญนั้นผู้วิจัยได้กำหนดพื้นที่ในแต่ละด้านเอาไว้ ดังนี้ พื้นที่ทางด้านซ้ายและพื้นที่ทางด้านขวากำหนดเป็น 35% ของพื้นที่ทั้งหมดในรูปภาพและพื้นที่ตรงกลางกำหนดเป็น 50% ของพื้นที่ทั้งหมดในรูปภาพ ดังรูปที่ 4.9 โดยกำหนดให้ระหว่างพื้นที่ของแต่ละด้านที่ทับซ้อนกันอยู่ให้ซ้อนทับกันด้านละ 10%



รูปที่ 4.9 พื้นที่ตามความสำคัญของแต่ละด้าน

#### 4.2.2 การดึงคุณลักษณะสำคัญของเสื้อผ้า (Clothing features extraction)

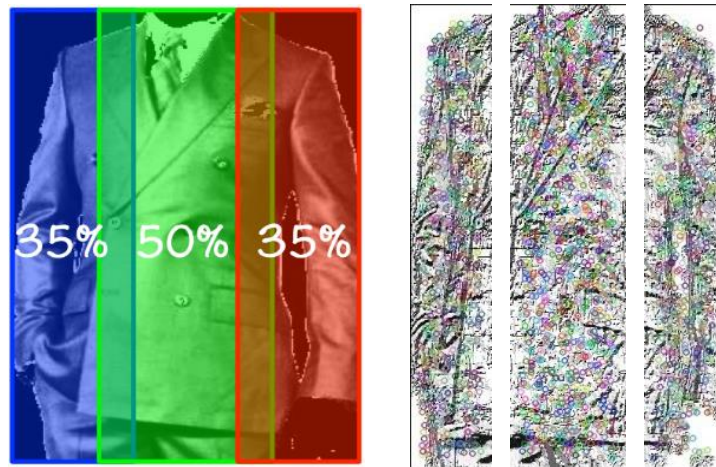
กระบวนการดึงคุณลักษณะสำคัญของเสื้อผ้า เพื่อนำไปใช้ในเทคนิคของรูปภาพของคุณลักษณะสำคัญ ในการเปรียบเทียบชนิดของเสื้อผ้านั้น ผู้วิจัยได้นำเสนอวิธีการใหม่ในการเพิ่มความถูกต้องของการรู้จำและแบ่งกลุ่มชนิดของเสื้อผ้า โดยใช้ Local Binary Patterns ร่วมกับ Speeded-Up Robust Features ด้วยเทคนิครูปภาพของคุณลักษณะสำคัญ

ผู้วิจัยได้ใช้ Local Binary Patterns ในการลดสัญญาณรบกวนต่างๆ ที่มาจากการเปลี่ยนแปลงของแสงหรือความเข้มของแสงและสัญญาณรบกวนที่ไม่มีรูปแบบเฉพาะ ดังรูปที่ 4.10 แสดงผลของภาพเสื้อผ้าหลังผ่านการคำนวณด้วย Local Binary Patterns จะเห็นได้ว่าภาพที่ได้จะมีลักษณะเป็นภาพขาว ดำ และแสดงโครงสร้างของเสื้อผ้าอย่างชัดเจน



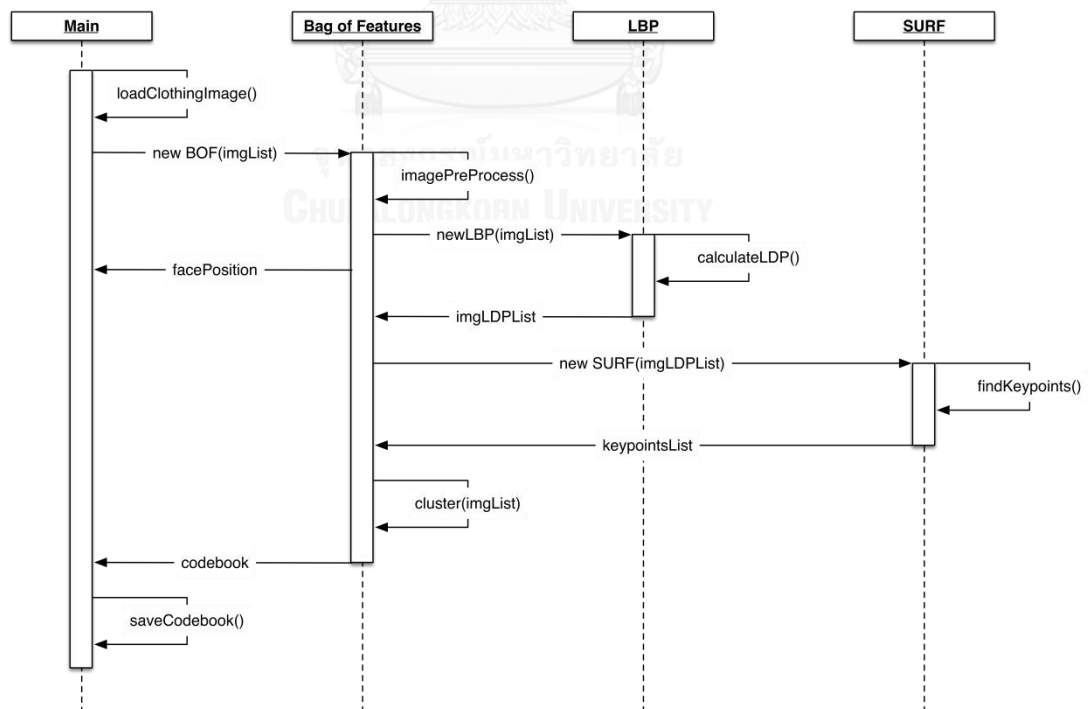
รูปที่ 4.10 ผลที่ได้จากการทำ Local Binary Patterns

ภาพที่ผ่านการลดสัญญาณรบกวนด้วยวิธีคำนวณผ่าน Local Binary Patterns จะถูกนำไปค้นหาจุดสนใจในขั้นตอนต่อไปด้วยวิธีการของ Speeded-Up Robust Features ในขั้นตอนนี้จุดสนใจจะถูกให้น้ำหนักแตกต่างกันตามขนาดพื้นที่ที่กำหนดไว้ในการจัดการข้อมูลนำเข้าก่อนการวิเคราะห์ โดยในพื้นที่ด้านซ้ายและพื้นที่ด้านขวาจะกำหนดน้ำหนักของจุดสนใจเป็น 100 จุดสนใจ ส่วนพื้นที่ด้านตรงกลางจะกำหนดน้ำหนักของจุดสนใจเป็น 300 จุดสนใจ ดังรูปที่ 4.11



รูปที่ 4.11 หน้าหนักของจุดสนใจในแต่ละพื้นที่

กระบวนการสร้างพจนานุกรมของจุดสนใจตามเทคนิคของคุณภาพของคุณลักษณะสำคัญ นั้นจะใช้จุดสนใจของเสื้อผ้ามาจัดเป็นกลุ่มจุดสนใจที่มีชนิดเดียวกันเอาไว้กลุ่มเดียวกัน โดยเปรียบเทียบกับชุดข้อมูลในตัวอย่างฝึกสอนทั้งหมดมาจัดกลุ่มจุดสนใจเหล่านั้นด้วยวิธีการจัดกลุ่มแบบ K-mean ซึ่งในวิทยานิพนธ์นี้ผู้วิจัยได้แบ่งจำนวนกลุ่มของ K-mean เอาไว้ทั้งหมด 1024 กลุ่ม [29] ซึ่งผู้วิจัยได้นำเสนอในรูปแบบของแผนภาพลำดับของกระบวนการดึงคุณลักษณะสำคัญ เพื่อนำไปใช้เป็นพจนานุกรมของจุดสนใจ ดังรูปที่ 4.12



รูปที่ 4.12 แผนภาพลำดับของกระบวนการสร้างคุณลักษณะสำคัญ



### 4.3 การสอนให้เครื่องเรียนรู้และการทดสอบ (Clothing Classification)

หลังจากผ่านขั้นตอนของการจัดกลุ่มแบบ K-mean เพื่อนำไปสร้างเป็นพจนานุกรมของจุดสนใจ (Codebook) จากนั้นภาพในชุดข้อมูลฝึกสอนที่นำมาใช้ในการสร้างพจนานุกรมนั้นจะถูกนำไปค้นหาจุดสนใจอีกรอบเพื่อใช้ในการสร้างคุณลักษณะสำคัญของจุดสนใจแล้วนำไปเปรียบเทียบกับคุณลักษณะสำคัญของจุดสนใจในพจนานุกรมว่าแต่ละคุณลักษณะสำคัญของจุดสนใจตกอยู่ในคลัสเตอร์ใดบ้าง ผลลัพธ์ที่ได้ออกมาเพื่อนำไปใช้ในการสอนระบบให้เรียนรู้รูปและทดสอบนั้นจะอยู่ในรูปแบบของ Histogram of Features คือ จะเป็นการนับจำนวนจุดสนใจทั้งหมดในรูปภาพว่าตกอยู่ในแต่ละคลัสเตอร์เป็นเท่าใด

ขั้นตอนการสอนให้ระบบเรียนรู้และทดสอบในวิทยานิพนธ์ฉบับนี้ผู้วิจัยได้นำเสนอ โดยใช้วิธีการเรียนรู้แบบซัพพอร์ตเวกเตอร์แมชชีนมาใช้ในการรู้จำชนิดของเสื้อผ้า ซึ่งในขั้นตอนการสอนให้ระบบเรียนรู้นั้นค่อนข้างใช้เวลานานเพราะเป็นการเรียนรู้ของข้อมูลในจำนวนมาก ซึ่งมีขนาดของข้อมูลแต่ละค่าเป็น 1024 เวกเตอร์ ดังนั้นผู้วิจัยได้ทำการแยกการออกแบบระบบออกเป็นสองส่วนของขั้นตอนนี้ คือ ส่วนของสอนให้ระบบเรียนรู้แล้วให้สร้างโมเดลออกมาเพื่อนำไปใช้ในการรู้จำได้โดยไม่ต้องส่งข้อมูลมาสอนใหม่ และขั้นตอนการทดสอบระบบ โดยนำโมเดลที่ได้จากขั้นตอนนี้ก่อนหน้าไปรู้จำชนิดของเสื้อผ้า

จากรูปที่ 4.13 แสดงการสร้างชุดคุณลักษณะสำคัญที่ใช้ในการฝึกสอนระบบให้เรียนรู้จากวิธีการของ Bag of Features คือ จะนับจำนวนของจุดสนใจในภาพแต่ละจุดว่าไปตกอยู่ในคลัสเตอร์ใดบ้างและนำไปสร้างเป็นคุณลักษณะสำคัญสำหรับการสอนการเรียนรู้ ซึ่งจะเห็นได้ว่าคุณลักษณะสำคัญที่ได้มาจะมีขนาดของเวกเตอร์เท่ากับขนาดของพจนานุกรมของภาพ ในวิทยานิพนธ์ได้กำหนดขนาดของพจนานุกรมของภาพไว้ที่ 1,024 ต่อหนึ่งภาพ ซึ่งการฝึกสอนให้ระบบเรียนรู้นั้น จะใช้ภาพของชุดข้อมูลฝึกสอนเท่านั้น โดยแยกออกจากภาพของชุดข้อมูลทดสอบระบบ ผลลัพธ์ที่ได้ออกมาจากการสอนให้ระบบเรียนรู้ คือ โมเดลของซัพพอร์ตเวกเตอร์แมชชีนแบบ linear และทำการบันทึกโมเดลนั้นลงฐานข้อมูล



รูปที่ 4.13 แสดงการควบคุมขั้นตอนการทำงานของระบบการสอน

#### 4.3.1 ขั้นตอนการเตรียมข้อมูลทดสอบ

จำนวนรูปภาพที่ใช้ในงานวิจัยนี้ทั้งหมด 1131 รูปภาพ โดยได้แบ่งเป็น 7 ชนิดด้วยกัน คือ รูปภาพเสื้อแจ๊คเก็ต เสื้อเชิ้ต เสื้อสูท เสื้อกันหนาว เสื้อยืด เสื้อโปโล และเสื้อกั๊ก ซึ่งใช้จำนวนรูปภาพเฉลี่ยชนิดละประมาณ 100-170 ภาพ ในการสอนให้ระบบเรียนรู้และ ใช้รูปภาพ 20 ภาพต่อหนึ่งชนิดสำหรับการทดสอบ โดยภาพที่นำมาใช้ในงานวิจัยนี้มาจากฐานข้อมูลของ Google imageNet [30] ฐานข้อมูลร้านขายเสื้อผ้าออนไลน์ และข้อมูลจากการถ่ายภาพด้วยกล้องดิจิทัลและกล้องเว็บแคม ซึ่งทำการเลือกภาพที่มีลักษณะเด่น คือ เป็นภาพใบหน้าตรง ภาพเต็มตัวหรือภาพครึ่งตัวร่างกายส่วนบน และไม่มีสิ่งขีดขวางบดบังเสื้อผ้าในการทดลองนี้ ดังแสดงในตารางที่ 4.1

ตารางที่ 4.1 จำนวนรูปภาพที่ใช้ในแต่ละกลุ่ม

	เสื้อแจ๊คเก็ต	เสื้อเชิ้ต	เสื้อยืด	เสื้อโปโล	เสื้อสูท	เสื้อกันหนาว	เสื้อกั๊ก
สอน	140	120	121	160	170	135	145
ทดสอบ	20	20	20	20	20	20	20



(ก) เสื้อแจ๊คเก็ต [31]



(ข) เสื้อเชิ้ต [32]



(ค) เสื้อยืด [33]



(ง) เสื้อโปโล [34]



(จ) เสื้อสูท [35]



(ฉ) เสื้อกันหนาว [36]



(ช) เสื้อกล้าม [37]

รูปที่ 4.14 ตัวอย่างรูปภาพทั้ง 7 ชนิด

#### 4.3.2 วิธีการวัดประสิทธิผลความถูกต้อง

วิธีการรู้จำชนิดของเสื้อผ้า นั้นได้ประเมินประสิทธิผลความถูกต้องของการรู้จำ โดยใช้ซอฟต์แวร์เวกเตอร์แมชชีนหลังจากการสอนให้ระบบเรียนรู้และทดสอบระบบ อีกทั้งเพื่อใช้เป็นมาตรฐานชี้วัดคุณภาพของการสอนให้ระบบเรียนรู้ได้ถูกต้อง ซึ่งในวิทยานิพนธ์ฉบับนี้ได้รวบรวมวิธีการประเมินประสิทธิผลทั้งหมด 2 วิธีการด้วยกัน ดังนี้

- 1) การคำนวณหาค่าความถูกต้องของข้อมูลทั้งหมด (Accuracy Score) [38] เป็นการเปรียบเทียบระหว่างข้อมูลที่นำเข้ามาทดสอบและข้อมูลของภาพในโมเดลว่ามีอัตราตรงตามเกณฑ์ของจำนวนความถูกต้องทั้งหมดหรือไม่ ดังสมการที่ (16)
- 2) การคำนวณหาค่าความแม่นยำ (Precision) ค่าความครบถ้วน (Recall) และ F-score [38] เป็นวิธีการวัดค่าความถูกต้องของข้อมูลที่สามารถทำนายได้ตรงตามความเป็นจริง โดยค่า Precision จะเป็นวิธีการคำนวณหาความเกาะกลุ่มของชุดข้อมูลที่ทำนาย ซึ่งค่า Precision ที่เหมาะสมควรมีความเกาะกลุ่มกัน ดังสมการที่ (17) ส่วน Recall เป็นการวัดชุดข้อมูลที่ทำนายถูกกับกลุ่มข้อมูลชุดนั้น ดังสมการที่ (18) และ F-score คือ สัดส่วนความถูกต้อง ซึ่งการคำนวณจะใช้ค่าของความแม่นยำและค่าความครบถ้วน ดังสมการที่ (19)

ตารางที่ 4.2 การวัดประสิทธิผลของข้อมูล

		ค่าจริงของข้อมูล	
ค่าที่ได้จากการทำนาย	TP (True Positive)	FP (False Positive)	
	FN (False Negative)	TN (True Negative)	

ตารางที่ 4.3 สูตรการคำนวณค่า Accuracy, Precision และ Recall

$$\text{Accuracy} = \frac{TP + FN}{TP + TN + FP + FN} \times 100 \quad (16)$$

$$\text{Precision} = \frac{TP}{TP + FP} \times 100 \quad (17)$$

$$Recall = \frac{TP}{TP + FN} \times 100 \quad (18)$$

$$F\ score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \times 100 \quad (19)$$

โดยที่ Accuracy คือ ค่าความถูกต้อง

TP คือ จำนวนข้อมูลที่ทายถูกที่เป็นกลุ่มบวก

FN คือ จำนวนข้อมูลที่ทายถูกที่เป็นกลุ่มลบ

FP คือ จำนวนข้อมูลที่ทายผิดที่เป็นกลุ่มบวก

TN คือ จำนวนข้อมูลที่ทายผิดที่เป็นกลุ่มลบ

- 3) K-fold Cross-validation เป็นวิธีการที่ใช้ในงานวิจัยเพื่อทดสอบความถูกต้องโดยวิธีการแบ่งข้อมูลทั้งหมด K กลุ่ม กลุ่มละเท่าๆ กันแล้วกำหนดให้กลุ่มที่หนึ่งเป็นตัวทดสอบที่เหลือทั้งหมดเป็นข้อมูลสอนให้ระบบเรียนรู้แล้วดูค่าความถูกต้องหลังจากนั้นให้ทำใหม่โดยกำหนดให้กลุ่มที่สองเป็นตัวทดสอบทำงานถึงตัวที่ K แล้วดูว่าค่าความถูกต้องทั้งหมดสามารถยอมรับได้หรือไม่ ตัวอย่างในตารางที่ 4.4 ได้กำหนดค่า k เท่ากับ 5

ตารางที่ 4.4 k-fold cross-validation

ชุดข้อมูล จำนวนครั้ง	1	2	3	4	5
1	ชุดทดสอบ	ชุดสอน	ชุดสอน	ชุดสอน	ชุดสอน
2	ชุดสอน	ชุดทดสอบ	ชุดสอน	ชุดสอน	ชุดสอน
3	ชุดสอน	ชุดสอน	ชุดทดสอบ	ชุดสอน	ชุดสอน
4	ชุดสอน	ชุดสอน	ชุดสอน	ชุดทดสอบ	ชุดสอน
5	ชุดสอน	ชุดสอน	ชุดสอน	ชุดสอน	ชุดทดสอบ

## บทที่ 5

### การทดลองและผลการวิจัย

ในบทนี้จะกล่าวถึงรายละเอียดของการทดลองตามวิธีต่างๆ และที่วิทยานิพนธ์ฉบับนี้ได้นำเสนอ หลังจากนั้นจะนำผลการทดสอบความถูกต้องไปทำการเปรียบเทียบกับวิธีการอื่นๆ คือ การค้นหาจุดสนใจโดยใช้ SIFT การค้นหาจุดสนใจโดยใช้ SURF การค้นหาจุดสนใจโดยใช้ SURF ร่วมกับ LBP การค้นหาจุดสนใจโดยใช้ SURF การค้นหาจุดสนใจโดยใช้ SURF ร่วมกับ LBP และวิธีที่วิทยานิพนธ์ฉบับนี้ได้นำเสนอ คือ การค้นหาจุดสนใจแบบถ่วงน้ำหนักโดยใช้ SURF ร่วมกับ LBP โดยอุปกรณ์ที่ใช้ในการทำงานวิจัย ประกอบไปด้วย

#### 1) เครื่องคอมพิวเตอร์

- Intel® Core™ 2 Duo CPU 2.27 GHz
- Memory 5.00 GB
- Hard Drive 80.0 GB 5400 rpm

#### 2) ซอฟต์แวร์

- Operating System Microsoft Windows 7 Professional
- IDE Qt Creator Version 5.4.2
- Language and GUI C++ OpenCV library

### 5.1 ผลการทดสอบความถูกต้องในแต่ละวิธี

ผลการทดสอบความถูกต้องในแต่ละวิธีได้ใช้จำนวนรูปภาพทดสอบทั้งหมด 140 ภาพ ซึ่งแบ่งออกเป็น 7 ชนิดเท่าๆ กันในตารางที่ 4.1 และได้ใช้การเรียนรู้ด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนแบบ linear kernel และทำการปรับค่าพารามิเตอร์ C เท่ากับ 312.50 และ gamma เท่ากับ 1.00 ซึ่งในวิทยานิพนธ์นี้ได้แสดงผลของวิธีการรู้จำชนิดของเสื้อผ้า ดังนี้

- การค้นหาจุดสนใจโดยใช้ SIFT

วิธีการค้นหาจุดสนใจโดยใช้ SIFT เป็นวิธีการที่ถูกใช้ในการดูความคล้ายของรูปภาพ ซึ่งผลการทดสอบความถูกต้องในการนำมาใช้เพื่อรู้จำชนิดของเสื้อผ้านั้นแสดงดังตารางที่ 5.1 แสดงผลของการทดสอบตามชนิดของเสื้อผ้า โดยในแนวนอนของตารางจะหมายถึงค่าที่ระบบทำนายออกมา

(Predicted) ส่วนในแนวตั้งจะหมายถึงค่าจริงที่ส่งให้ระบบทำนาย (Actual) และค่าความถูกต้องเท่ากับ 48.57% ของตัวอย่างที่นำมาทดสอบทั้งหมด

ตารางที่ 5.1 Confusion matrix ของแต่ละชนิดเสื้อผ้าด้วยวิธี SIFT

	แจ๊คเก็ต	เชิ้ต	สูท	กันหนาว	ยืด	โปโล	กล้ามเนื้อ
แจ๊คเก็ต	16	1	0	0	0	2	1
เชิ้ต	1	4	1	1	0	2	0
สูท	0	8	17	1	2	0	0
กันหนาว	0	1	0	10	1	4	4
ยืด	2	0	0	2	4	2	2
โปโล	1	2	2	0	8	8	4
กล้ามเนื้อ	0	4	0	6	5	2	9

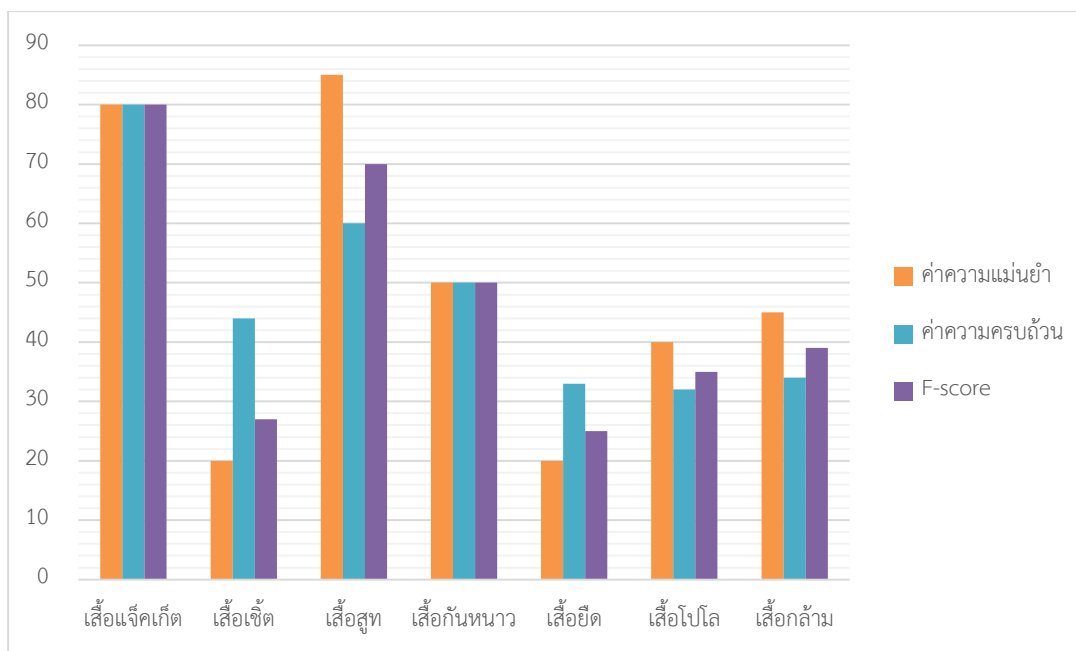
ตารางที่ 5.2 แสดงรายละเอียดของค่าความแม่นยำ, ค่าความครบถ้วนและค่า F-score โดยแบ่งย่อยกลุ่มการทดสอบออกเป็นตามชนิดของเสื้อผ้า ซึ่งจะเห็นได้ว่าค่าสูงสุดของกลุ่มทดสอบ คือ เสื้อผ้าชนิดเสื้อสูท ที่มีค่าความแม่นยำเท่ากับ 85% ค่าความครบถ้วนเท่ากับ 60% และค่า F-score เท่ากับ 70%

ตารางที่ 5.2 รายละเอียดของค่าความแม่นยำ ค่าความครบถ้วนและค่า F-score ด้วยวิธี SIFT

	ค่าความแม่นยำ	ค่าความครบถ้วน	F-score
เสื้อแจ๊คเก็ต	80.00%	80.00%	80.00%
เสื้อเชิ้ต	20.00%	44.44%	27.59%
เสื้อสูท	85.00%	60.71%	70.83%
เสื้อกันหนาว	50.00%	50.00%	50.00%
เสื้อยืด	20.00%	33.33%	25.00%
เสื้อโปโล	40.00%	32.00%	35.56%
เสื้อกล้ามเนื้อ	45.00%	34.62%	39.13%
เฉลี่ยรวมทั้งหมด	48.57%	47.87%	46.87%

รูปที่ 5.1 แสดงกราฟแผนภูมิแนวตั้งของค่าความแม่นยำและค่าความครบถ้วนในแต่ละกลุ่มการทดสอบ โดยในแนวตั้ง คือ อันดับคะแนน ซึ่งคิดเป็นเปอร์เซ็นต์และในแนวนอน คือ ชนิดของเสื้อผ้าทั้งหมด 7 ชนิด





รูปที่ 5.1 กราฟแสดงค่าของความแม่นยำและความถูกต้องในแต่ละกลุ่มการทดสอบด้วยวิธี SIFT

- การค้นหาจุดสนใจโดยใช้ SURF

วิธีการค้นหาจุดสนใจด้วย SURF นั้นเป็นวิธีที่พัฒนาต่อจาก SIFT ในเรื่องของความเร็วและความถูกต้อง ดังนั้นในงานวิจัยนี้จึงได้นำวิธีการ SURF เพื่อรู้จำชนิดของเสื้อผ้า ซึ่งผลการทดสอบความถูกต้องแสดงในตารางที่ 5.3 แสดงผลของการทดสอบตามชนิดของเสื้อผ้า โดยในแนวนอนของตารางจะหมายถึงค่าที่ระบบทำนายออกมา (Predicted) ส่วนในแนวตั้งจะหมายถึงค่าจริงที่ส่งให้ระบบทำนาย (Actual) และค่าความถูกต้องเท่ากับ 57.86% ของตัวอย่างที่นำมาทดสอบทั้งหมด

ตารางที่ 5.3 Confusion matrix ของแต่ละชนิดเสื้อผ้าด้วยวิธี SURF

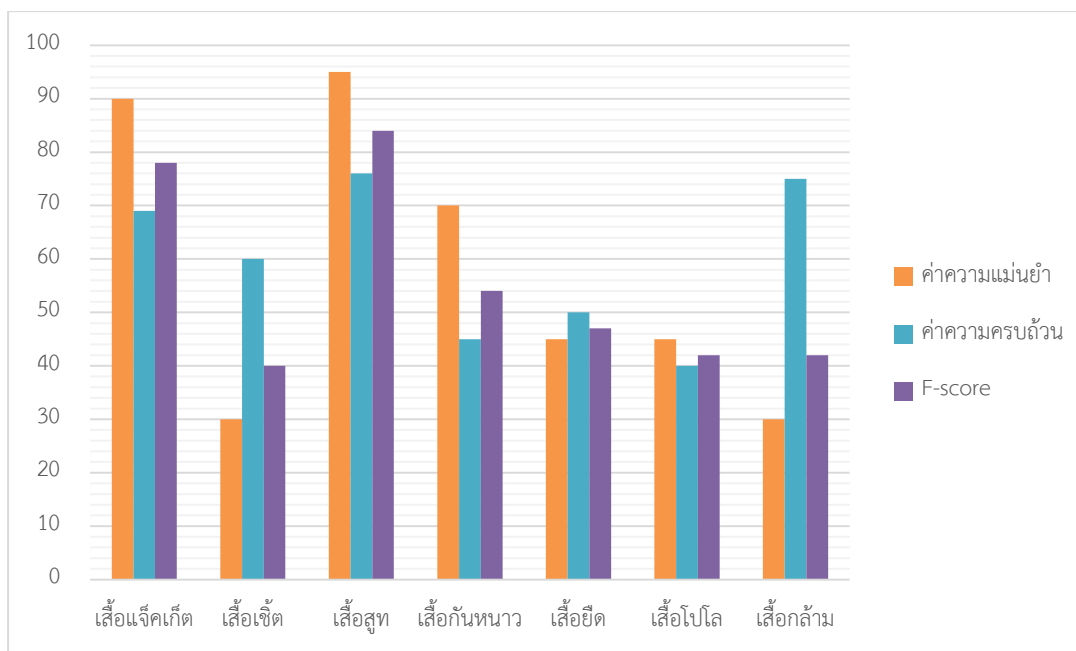
	แจ็กเก็ต	เชิ้ต	สูท	ก้นหนาว	ยืด	โปโล	กล้าม
แจ็กเก็ต	18	5	1	1	0	0	1
เชิ้ต	0	6	0	4	0	0	0
สูท	0	5	19	0	0	0	1
ก้นหนาว	0	4	0	14	1	5	7
ยืด	2	0	0	1	9	4	2
โปโล	0	0	0	0	10	9	3
กล้าม	0	0	0	0	0	2	6

ตารางที่ 5.4 แสดงรายละเอียดของค่าความแม่นยำ, ค่าความครบถ้วนและค่า F-score โดยแบ่งย่อยกลุ่มการทดสอบออกเป็นตามชนิดของเสื้อผ้า ซึ่งจะเห็นได้ว่าค่าสูงสุดของกลุ่มทดสอบ คือ เสื้อผ้าชนิดเสื้อสูท ที่มีค่าความแม่นยำเท่ากับ 95% ค่าความครบถ้วนเท่ากับ 76% และค่า F-score เท่ากับ 84%

ตารางที่ 5.4 รายละเอียดของค่าความแม่นยำ ค่าความครบถ้วนและค่า F-score ด้วยวิธี SURF

	ค่าความแม่นยำ	ค่าความครบถ้วน	F-score
เสื้อแจ็กเก็ต	90.00%	69.23%	78.26%
เสื้อเชิ้ต	30.00%	60.00%	40.00%
เสื้อสูท	95.00%	76.00%	84.44%
เสื้อกันหนาว	70.00%	45.16%	54.90%
เสื้อยืด	45.00%	50.00%	47.37%
เสื้อโปโล	45.00%	40.91%	42.86%
เสื้อกล้าม	30.00%	75.00%	42.86%
เฉลี่ยรวมทั้งหมด	57.86%	59.47%	55.81%

รูปที่ 5.2 แสดงกราฟแผนภูมิแนวตั้งของค่าความแม่นยำและค่าความครบถ้วนในแต่ละกลุ่มการทดสอบ โดยในแนวตั้ง คือ อันดับคะแนน ซึ่งคิดเป็นเปอร์เซ็นต์และในแนวนอน คือ ชนิดของเสื้อผ้า ทั้งหมด 7 ชนิด



รูปที่ 5.2 กราฟแสดงค่าของความแม่นยำและความถูกต้องในแต่ละกลุ่มการทดสอบด้วยวิธี SURF

- การค้นหาจุดสนใจโดยใช้ SIFT ร่วมกับ LBP

การใช้ Local texture เข้ามาดำเนินการกับรูปภาพก่อนที่จะหาจุดสนใจ ผู้วิจัยคิดว่าสามารถเพิ่มประสิทธิภาพให้กับการเรียนรู้ได้เพื่อนำมาใช้รู้จำชนิดของเสื้อผ้า ตารางที่ 5.5 แสดงผลของการทดสอบตามชนิดของเสื้อผ้า โดยในแนวนอนของตารางจะหมายถึงค่าที่ระบบทำนายออกมา (Predicted) ส่วนในแนวตั้งจะหมายถึงค่าจริงที่ส่งให้ระบบทำนาย (Actual) และค่าความถูกต้องเท่ากับ 61.35% ของตัวอย่างที่นำมาทดสอบทั้งหมด

ตารางที่ 5.5 Confusion matrix ของแต่ละชนิดเสื้อผ้าด้วยวิธี SIFT ร่วมกับ LBP

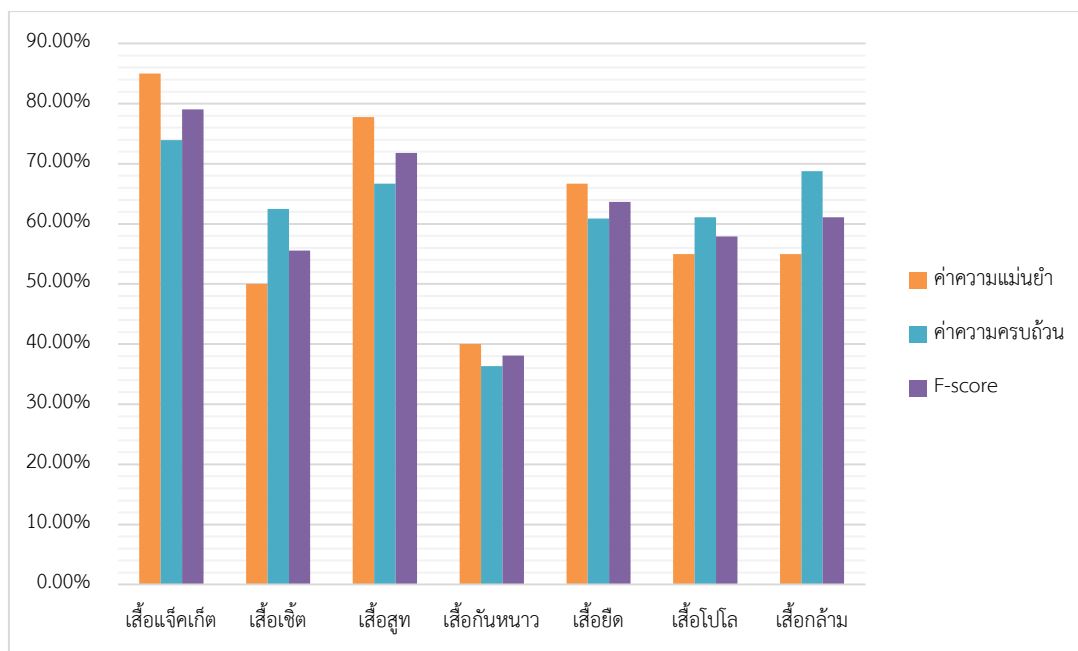
	แจ็กเก็ต	เชิ้ต	สูท	กัณหาว	ยืด	โปโล	กัล้าม
แจ็กเก็ต	17	2	0	2	1	1	0
เชิ้ต	1	10	1	3	0	0	1
สูท	2	2	14	2	0	1	0
กัณหาว	0	4	2	8	1	1	6
ยืด	0	1	0	3	14	4	1
โปโล	0	0	1	1	4	11	1
กัล้าม	0	1	0	1	1	2	11

ตารางที่ 5.6 แสดงรายละเอียดของค่าความแม่นยำ, ค่าความครบถ้วนและค่า F-score โดยแบ่งย่อยกลุ่มการทดสอบออกเป็นตามชนิดของเสื้อผ้า ซึ่งจะเห็นได้ว่าค่าสูงสุดของกลุ่มทดสอบ คือ เสื้อผ้าชนิดเสื้อสูท ที่มีค่าความแม่นยำเท่ากับ 77% ค่าความครบถ้วนเท่ากับ 66% และค่า F-score เท่ากับ 71%

ตารางที่ 5.6 รายละเอียดของค่าความแม่นยำ ค่าความครบถ้วนและค่า F-score ด้วยวิธี SIFT ร่วมกับ LBP

	ค่าความแม่นยำ	ค่าความครบถ้วน	F-score
เสื้อแจ็คเก็ต	85.00%	73.91%	79.07%
เสื้อเชิ้ต	50.00%	62.50%	55.56%
เสื้อสูท	77.78%	66.67%	71.79%
เสื้อกันหนาว	40.00%	36.36%	38.10%
เสื้อยืด	66.67%	60.87%	63.64%
เสื้อโปโล	55.00%	61.11%	57.89%
เสื้อกล้าม	55.00%	68.75%	61.11%
เฉลี่ยรวมทั้งหมด	61.35%	61.45%	61.02%

รูปที่ 5.3 แสดงกราฟแผนภูมิแนวตั้งของค่าความแม่นยำและค่าความครบถ้วนในแต่ละกลุ่มการทดสอบ โดยในแนวตั้ง คือ อันดับคะแนน ซึ่งคิดเป็นเปอร์เซ็นต์และในแนวนอน คือ ชนิดของเสื้อผ้าทั้งหมด 7 ชนิด



รูปที่ 5.3 กราฟแสดงค่าของความแม่นยำและความถูกต้องในแต่ละกลุ่มการทดสอบด้วยวิธี SIFT ร่วมกับ LBP

- การค้นหาจุดสนใจโดยใช้ SURF ร่วมกับ LDP

ตารางที่ 5.7 แสดงผลของการทดสอบตามชนิดของเสื้อผ้า โดยในแนวนอนของตารางจะหมายถึงค่าที่ระบบทำนายออกมา (Predicted) ส่วนในแนวตั้งจะหมายถึงค่าจริงที่ส่งให้ระบบทำนาย (Actual) และค่าความถูกต้องเท่ากับ 62.14% ของตัวอย่างที่นำมาทดสอบทั้งหมด

ตารางที่ 5.7 Confusion matrix ของแต่ละชนิดเสื้อผ้าด้วยวิธี SURF ร่วมกับ LDP

	แฉกเกิด	เชื้อ	สู้	ก้นหนาว	ยืด	โปโล	กล้าม
แฉกเกิด	18	4	0	1	0	0	1
เชื้อ	0	8	0	8	1	0	1
สู้	0	5	20	0	0	0	1
ก้นหนาว	0	3	0	7	1	3	3
ยืด	2	0	0	1	12	3	0
โปโล	0	0	0	2	6	9	1
กล้าม	0	0	0	1	0	5	13

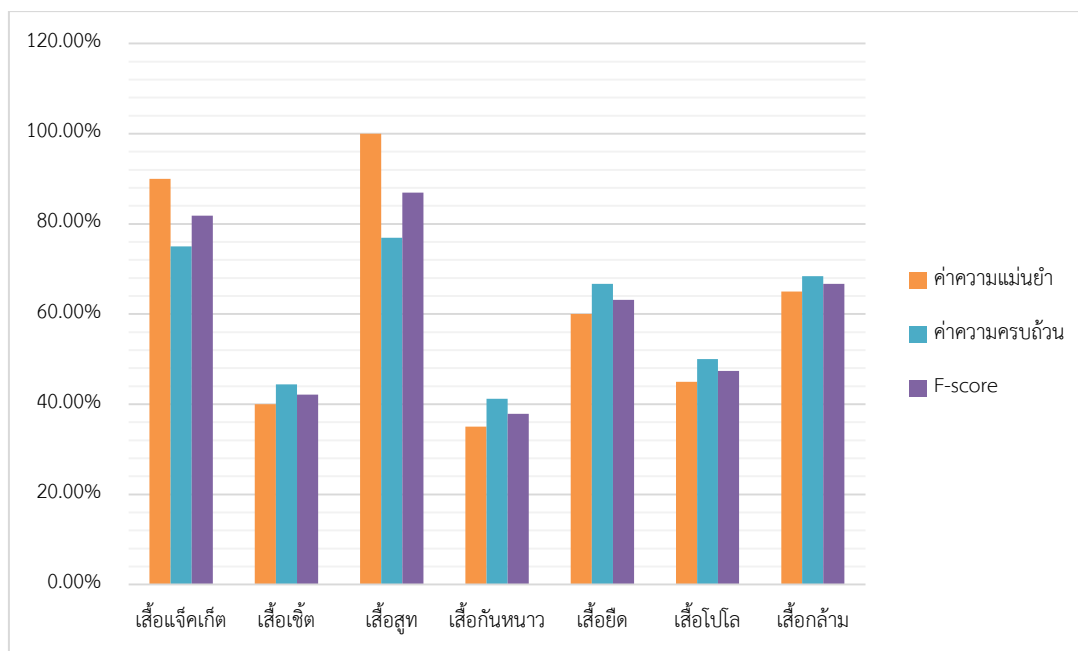
ตารางที่ 5.8 แสดงรายละเอียดของค่าความแม่นยำ, ค่าความครบถ้วนและค่า F-score โดยแบ่งย่อยกลุ่มการทดสอบออกเป็นตามชนิดของเสื้อผ้า ซึ่งจะเห็นได้ว่าค่าสูงสุดของกลุ่มทดสอบ คือ

เสื้อผ้าชนิดเสื้อสูท ที่มีค่าความแม่นยำเท่ากับ 100% ค่าความครบถ้วนเท่ากับ 76% และค่า F-score เท่ากับ 86%

ตารางที่ 5.8 รายละเอียดของค่าความแม่นยำ ค่าความครบถ้วนและค่า F-score ด้วยวิธี SURF ร่วมกับ LDP

	ค่าความแม่นยำ	ค่าความครบถ้วน	F-score
เสื้อแจ็กเก็ต	90.00%	75.00%	81.82%
เสื้อเชิ้ต	40.00%	44.44%	42.11%
เสื้อสูท	100.00%	76.92%	86.96%
เสื้อกันหนาว	35.00%	41.18%	37.84%
เสื้อยืด	60.00%	66.67%	63.16%
เสื้อโปโล	45.00%	50.00%	47.37%
เสื้อกล้าม	65.00%	68.42%	66.67%
เฉลี่ยรวมทั้งหมด	62.14%	60.38%	60.84%

รูปที่ 5.4 แสดงกราฟแผนภูมิแนวตั้งของค่าความแม่นยำและค่าความครบถ้วนในแต่ละกลุ่มการทดสอบ โดยในแนวตั้ง คือ อันดับคะแนน ซึ่งคิดเป็นเปอร์เซ็นต์และในแนวนอน คือ ชนิดของเสื้อผ้าทั้งหมด 7 ชนิด



รูปที่ 5.4 กราฟแสดงค่าของความแม่นยำและความถูกต้องในแต่ละกลุ่มการทดสอบด้วยวิธี SURF ร่วมกับ LDP

- การค้นหาจุดสนใจแบบถ่วงน้ำหนักโดยใช้ SURF ร่วมกับ LBP

วิทยานิพนธ์นี้ได้นำเสนอวิธีจำชนิดของเสื้อผ้า คือ ค้นหาจุดสนใจแบบถ่วงน้ำหนักโดยใช้ SURF ร่วมกับ LBP ด้วยเทคนิคคุณภาพของคุณลักษณะสำคัญ ตารางที่ 5.9 แสดงผลของการทดสอบตามชนิดของเสื้อผ้า โดยในแนวนอนของตารางจะหมายถึงค่าที่ระบบทำนายออกมา (Predicted) ส่วนในแนวตั้งจะหมายถึงค่าจริงที่ส่งให้ระบบทำนาย (Actual) ซึ่งในวิทยานิพนธ์นี้ประสบความสำเร็จได้ค่าความถูกต้องเท่ากับ 73.57% ของตัวอย่างที่นำมาทดสอบทั้งหมด

ตารางที่ 5.9 Confusion matrix ของแต่ละชนิดเสื้อผ้าด้วยวิธีที่นำเสนอ

	แฉกเกิด	เช็ด	สู้ท	ก้นหนาว	ยืด	โปโล	กล้ามเนื้อ
แฉกเกิด	18	0	2	0	0	0	0
เช็ด	2	10	2	6	0	0	0
สู้ท	0	1	19	0	0	0	0
ก้นหนาว	1	5	0	12	2	0	0
ยืด	0	0	0	0	18	2	0
โปโล	0	0	0	2	3	11	4
กล้ามเนื้อ	1	1	0	2	0	1	15

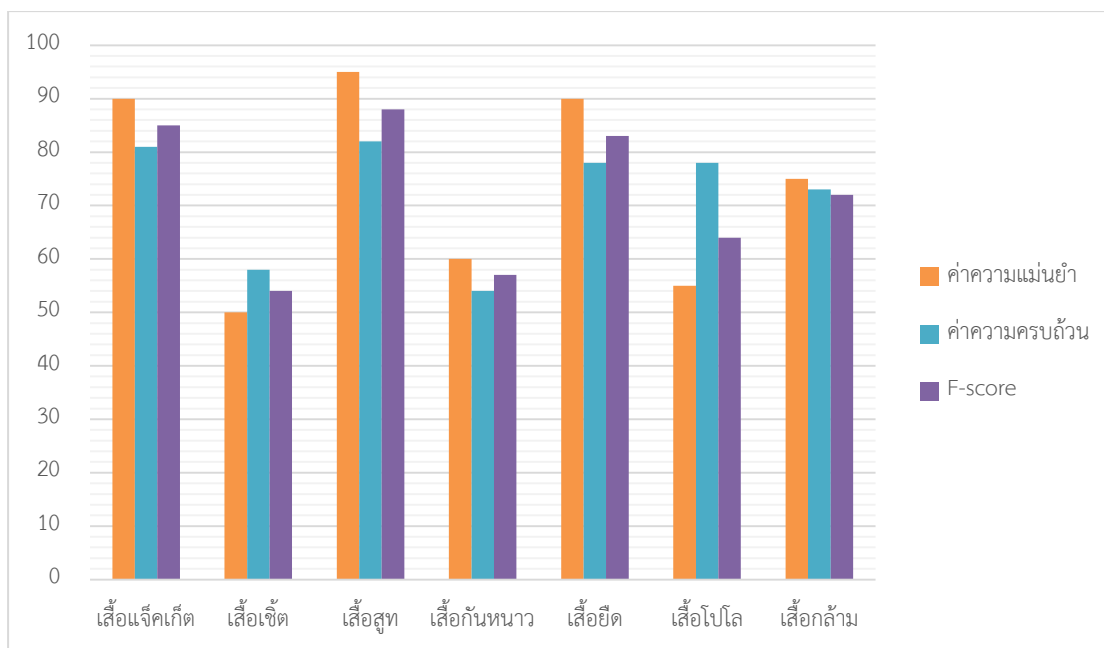
ตารางที่ 5.10 แสดงรายละเอียดของค่าความแม่นยำ, ค่าความครบถ้วนและค่า F-score โดยแบ่งย่อยกลุ่มการทดสอบออกเป็นตามชนิดของเสื้อผ้า ซึ่งจะเห็นได้ว่าค่าสูงสุดของกลุ่มทดสอบ คือ เสื้อผ้าชนิดเสื้อสูท ที่มีค่าความแม่นยำเท่ากับ 95% ค่าความครบถ้วนเท่ากับ 82% และค่า F-score เท่ากับ 88%

ตารางที่ 5.10 รายละเอียดของค่าความแม่นยำ ค่าความครบถ้วนและค่า F-score ด้วยวิธีที่นำเสนอ

	ค่าความแม่นยำ	ค่าความครบถ้วน	F-score
เสื้อแจ็คเก็ต	90.00%	81.82%	85.71%
เสื้อเชิ้ต	50.00%	58.82%	54.05%
เสื้อสูท	95.00%	82.61%	88.37%
เสื้อกันหนาว	60.00%	54.55%	57.14%
เสื้อยืด	90.00%	78.26%	83.72%
เสื้อโปโล	55.00%	78.57%	64.71%
เสื้อกล้าม	75.00%	78.95%	76.92%
เฉลี่ยรวมทั้งหมด	73.57%	73.37%	72.95%

รูปที่ 5.5 แสดงกราฟแผนภูมิแนวตั้งของค่าความแม่นยำและค่าความครบถ้วนในแต่ละกลุ่มการทดสอบ โดยในแนวตั้ง คือ อันดับคะแนน ซึ่งคิดเป็นเปอร์เซ็นต์และในแนวนอน คือ ชนิดของเสื้อผ้าทั้งหมด 7 ชนิด





รูปที่ 5.5 กราฟแสดงค่าของความแม่นยำและความถูกต้องในแต่ละกลุ่มการทดสอบด้วยวิธีที่นำเสนอ

## 5.2 ผลการเปรียบเทียบกับวิธีอื่นๆ

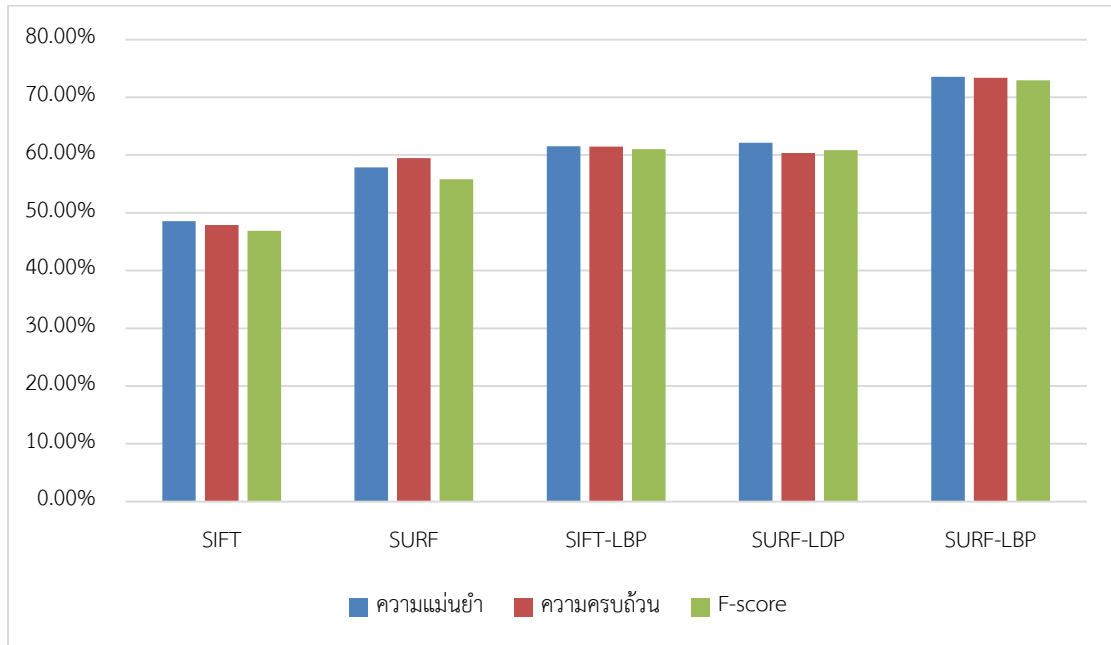
ในส่วนนี้จะเป็นการเปรียบเทียบวิธีการที่ได้นำเสนอมากับวิธีการก่อนหน้านี้ทั้งหมด 4 วิธีการ คือ การค้นหาจุดสนใจโดยใช้ SIFT การค้นหาจุดสนใจโดยใช้ SURF การค้นหาจุดสนใจโดยใช้ SURF ร่วมกับ LDP การค้นหาจุดสนใจโดยใช้ SURF การค้นหาจุดสนใจโดยใช้ SURF ร่วมกับ LDP

ตารางที่ 5.11 แสดงการเปรียบเทียบโดยรวมด้วยค่าความแม่นยำ ค่าความครบถ้วนและค่า F-score ซึ่งจะพบได้ว่าตามที่คุณวิจัยได้นำเสนอมีประสิทธิภาพผลความถูกต้องเพิ่มมากขึ้นเมื่อนำไปเปรียบเทียบกับวิธีการก่อนๆ

ตารางที่ 5.11 เปรียบเทียบวิธีต่างๆ ที่ได้ทดลองด้วยค่าความแม่นยำ ค่าความครบถ้วนและค่า F-score

	SIFT	SURF	SIFT-LBP	SURF-LDP	SURF-LBP
ความแม่นยำ	48.57%	57.86%	61.53%	62.14%	73.57%
ความครบถ้วน	47.87%	59.47%	61.45%	60.38%	73.37%
F-score	46.87%	55.81	61.02%	60.84%	72.95%

รูปที่ 5.6 แสดงกราฟการเปรียบเทียบวิธีการต่างๆ ที่ได้ทดลองรวมทั้งวิธีที่ผู้วิจัยได้นำเสนอ โดยแบ่งผลความถูกต้องตามชนิดของเสื้อผ้า วิธีการที่ผู้วิจัยได้นำเสนอนั้นเมื่อเปรียบเทียบประสิทธิภาพตามชนิดของเสื้อผ้าจะเห็นได้ว่ามีประสิทธิภาพในการรู้จำเป็นที่น่าพอใจในทุกๆ กลุ่มข้อมูล



รูปที่ 5.6 กราฟการเปรียบเทียบในแต่ละวิธีที่ได้ทดลอง โดยแบ่งผลความถูกต้องตามชนิดของเสื้อผ้า

## บทที่ 6

### สรุปผลการวิจัยและข้อเสนอแนะ

#### 6.1 สรุปผลการวิจัย

เป้าหมายของวิทยานิพนธ์ฉบับนี้ คือ ปรับปรุงประสิทธิภาพในการรู้จำชนิดของเสื้อผ้าให้มีความถูกต้องมากขึ้น ซึ่งผู้วิจัยได้นำเสนอวิธีการในการรู้จำชนิดของเสื้อผ้า โดยการใช้การหาจุดสนใจแบบถ่วงน้ำหนักด้วย Speeded-Up Robust Feature (SURF) ร่วมกับ Local Binary Patterns (LBP) และเทคนิคคุณภาพของคุณลักษณะสำคัญ ซึ่งมีแนวคิดว่า การถ่วงน้ำหนักของจุดสนใจไปยังพื้นที่ที่คิดว่าจะมีความสำคัญมากจะสามารถเพิ่มประสิทธิภาพของความถูกต้องมากขึ้น

ขั้นตอนการทำงานของวิธีการที่ได้นำเสนอนั้นได้อาศัยหลักการของการประมวลผลภาพเป็นส่วน เช่น การแบ่งส่วนของรูปภาพให้สามารถตัดส่วนที่เป็นฉากหลังออกได้ หรือแม้แต่การแบ่งส่วนของสีผิวมนุษย์ เป็นต้น ซึ่งหลังจากได้รูปภาพที่คิดว่าจะเป็นส่วนของเสื้อผ้ามาแล้วมาทำการดึงคุณลักษณะสำคัญเพื่อใช้ในการทดสอบ

การดึงคุณลักษณะสำคัญผู้วิจัยได้นำเสนอว่าให้ใช้ LBP ในการจัดการเรื่องของสัญญาณรบกวนต่างๆ เช่น การเปลี่ยนแปลงของแสง เป็นต้น แล้วจึงนำไปค้นหาจุดสนใจด้วย SURF การหาจุดสนใจด้วย SURF นั้นผู้วิจัยได้แบ่งการค้นหาจุดสนใจ โดยการแบ่งภาพออกเป็นสามส่วน คือ ด้านซ้าย ตรงกลางและด้านขวา ซึ่งส่วนที่ผู้วิจัยคิดว่าเป็นส่วนที่สำคัญ คือ ส่วนตรงกลางจึงได้ใส่น้ำหนักของการค้นหาจุดในบริเวณดังกล่าวมากกว่าด้านอื่นๆ ประมาณสามเท่า

ผลการทดลองพบว่าวิธีการที่ได้นำเสนอนั้นสามารถเพิ่มประสิทธิภาพของความถูกต้องได้ดีกว่าวิธีการค้นหาจุดสนใจที่นำมาเปรียบเทียบ โดยผลของความแม่นยำของสิ่งที่ได้นำเสนอนั้นได้เท่ากับ 73.75%

#### 6.2 ข้อเสนอแนะ

ในงานวิจัยฉบับนี้ยังมีส่วนที่ผิดพลาดอยู่เป็นจำนวนมากทั้งในเรื่องของการค้นหาจุดสนใจพื้นที่ที่ทำการค้นหาจุดสนใจ และขั้นตอนการเรียนรู้ ซึ่งได้แบ่งข้อผิดพลาดที่อาจจะเกิดขึ้น ได้ดังนี้

- การหาจุดสนใจยังไม่สามารถกำหนดจำนวนของจุดสนใจได้ว่าควรเป็นเท่าไร ซึ่งในการวิจัยได้ทดลองและพบว่าจำนวนการค้นหาจุดสนใจที่เหมาะสมกับวิทยานิพนธ์ฉบับนี้อยู่ที่ 100-300 จุดสนใจ

- พื้นที่ที่ทำการค้นหาจุดสนใจได้แบ่งออกเป็นสามส่วนหลักๆ ซึ่งการหาคุณลักษณะสำคัญที่ดีนั้นควรมีการกำหนดในหลากหลายรูปแบบ เช่น ทำเป็นตารางกริด เป็นต้น
- การแบ่งส่วนของฉากหลังและฉากหน้ายังคงพบความผิดพลาดของการแบ่งส่วนอยู่เป็นจำนวนมาก ซึ่งถ้าสีของเสื้อผ้าตรงกับสีของฉากหลังจะไม่สามารถแบ่งส่วนได้ถูกต้อง
- การเรียนรู้ของเครื่องจะพบว่าเป็นการเรียนรู้เพียงชนิดเดียว คือ ซัพพอร์ตเวกเตอร์แมชชีน ซึ่งคาดว่าชนิดของการเรียนรู้ในวิธีการอื่นๆ สามารถเพิ่มประสิทธิผลได้ด้วยเช่นกัน



## รายการอ้างอิง

1. Schouten., T. *Color models*. 2003 [cited 2015 June, 29]; Available from: <http://www.cs.ru.nl/~ths/rt2/col/h2/2fundENG.html#2.2.4>.
2. Kennelly, P.J. *HSV color model*. 2007 [cited 2015 June, 29]; Available from: <http://www.esri.com/news/arcuser/0807/nongeo.html>.
3. haltair. *Histogram equalization (part I)*. 2010; Available from: <https://haltair.wordpress.com/2010/07/17/histogram-equalization-part-i/>.
4. haltair. *Histogram equalization (part II)*. 2010; Available from: <https://haltair.wordpress.com/2010/07/17/histogram-equalization-part-ii/>.
5. Ziyae, T. *Haar wavelet*. 2013 [cited 2015 July, 14]; Available from: <http://dsp.stackexchange.com/questions/9964/how-to-remove-salt-pepper-noise-from-speech-signal>.
6. Viola, P. and M.J. Jones, *Robust Real-Time Face Detection*. *Int. J. Comput. Vision*, 2004. **57**(2): p. 137-154.
7. Ng, A. *CS229 Lecture notes: Support Vector Machines*. Machine Learning Course Materials 2014 [cited 2015 June, 6]; Available from: <http://cs229.stanford.edu/materials.html>.
8. Bradski, G. *The OpenCV Library*. OpenCV is an open-source, computer-vision library for extracting and processing meaningful data from images 2000 [cited 2015 June, 29]; Available from: <http://www.drdobbs.com/open-source/the-opencv-library/184404319>.
9. Wang, X. and T. Zhang, *Clothes search in consumer photos via color matching and attribute learning*, in *Proceedings of the 19th ACM international conference on Multimedia*. 2011, ACM: Scottsdale, Arizona, USA. p. 1353-1356.
10. Chiao-Meng, H., et al. *A sparse-coding based approach to clothing image retrieval*. in *Intelligent Signal Processing and Communications Systems (ISPACS), 2012 International Symposium on*. 2012.

11. Tianfa, W., et al. *Clothing image retrieval using color moment*. in *Computer Science and Network Technology (ICCSNT), 2013 3rd International Conference on*. 2013.
12. Miura, S., T. Yamasaki, and K. Aizawa. *SNAPPER: Fashion Coordinate Image Retrieval System*. in *Signal-Image Technology & Internet-Based Systems (SITIS), 2013 International Conference on*. 2013.
13. Mizuochi, M., A. Kanezaki, and T. Harada, *Clothing Retrieval Based on Local Similarity with Multiple Images*, in *Proceedings of the ACM International Conference on Multimedia*. 2014, ACM: Orlando, Florida, USA. p. 1165-1168.
14. Ahonen, T., A. Hadid, and M. Pietikainen, *Face Description with Local Binary Patterns: Application to Face Recognition*. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 2006. **28**(12): p. 2037-2041.
15. Kabir, M.H., T. Jabid, and O. Chae, *A Local Directional Pattern Variance (LDPv) Based Face Descriptor for Human Facial Expression Recognition*, in *Proceedings of the 2010 7th IEEE International Conference on Advanced Video and Signal Based Surveillance*. 2010, IEEE Computer Society. p. 526-532.
16. Kim, D.-J., S.-H. Lee, and M.-K. Sohn, *Face Recognition via Local Directional Pattern*. *International Journal of Security and Its Applications*, March 2013. **7**(2).
17. Lowe, D.G., *Distinctive Image Features from Scale-Invariant Keypoints*. *Int. J. Comput. Vision*, 2004. **60**(2): p. 91-110.
18. Sinha, U. *SIFT: Introduction*. *Computer Vision 2010* [cited 2015 July, 14]; Available from: <http://www.aishack.in/tutorials/sift-scale-invariant-feature-transform-keypoints/>.
19. Bay, H., T. Tuytelaars, and L. Van Gool, *SURF: Speeded Up Robust Features*, in *Computer Vision – ECCV 2006*, A. Leonardis, H. Bischof, and A. Pinz, Editors. 2006, Springer Berlin Heidelberg. p. 404-417.
20. Souza, C.d. *Haar-feature Object Detection in C#*. *integral image example 2014* [cited 2015 June, 29]; Available from:

<http://www.codeproject.com/Articles/441226/Haar-feature-Object-Detection-in-Csharp>.

21. Evans, C., *Notes on the OpenSURF Library*, in *CSTR-09-001*. January 2009: University of Bristol.
22. Lazebnik, S. *Bag-of-features models*. COMP 776: Computer Vision 2009 [cited 2015 July, 14]; Available from: [http://www.cs.unc.edu/~lazebnik/spring09/lec18\\_bag\\_of\\_features.pdf](http://www.cs.unc.edu/~lazebnik/spring09/lec18_bag_of_features.pdf).
23. O'Hara, S. and B.A. Draper, *Introduction to the Bag of Features Paradigm for Image Classification and Retrieval*. Computing Research Repository (CoRR), 2011.
24. Jusko, D.A. *Human Figure Drawing Proportions*. 2014; Available from: <http://www.realcolorwheel.com/human.htm>.
25. City-Suits.co.uk. *Navy Stripe Somerton Suit*. [cited 2015 July, 14]; Available from: <http://www.city-suits.co.uk/navy-stripe-somerton-suit.html>.
26. Rother, C., V. Kolmogorov, and A. Blake, "GrabCut": *interactive foreground extraction using iterated graph cuts*, in *ACM SIGGRAPH 2004 Papers*. 2004, ACM: Los Angeles, California. p. 309-314.
27. Blake, C.R.V.K.A., *GrabCut Interactive Foreground Extraction using Iterated Graph Cuts*. 2009.
28. Albiol, A., L. Torres, and E.J. Delp. *Optimum color spaces for skin detection*. in *Image Processing, 2001. Proceedings. 2001 International Conference on*. 2001.
29. Bossard, L., et al., *Apparel classification with style*, in *Proceedings of the 11th Asian conference on Computer Vision - Volume Part IV*. 2013, Springer-Verlag: Daejeon, Korea. p. 321-335.
30. Jia, D., et al. *ImageNet: A large-scale hierarchical image database*. in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. 2009.
31. Forever21. *Faux Leather Bomber Jacket*. [cited 2015 July, 7]; Available from: [http://www.forever21.com/UK/Product/Product.aspx?BR=F21&Category=outer\\_wear\\_faux-leather-jackets&ProductID=2000057500&VariantID=](http://www.forever21.com/UK/Product/Product.aspx?BR=F21&Category=outer_wear_faux-leather-jackets&ProductID=2000057500&VariantID=).

32. Forever21. *Colorblocked Collar Shirt*. [cited 2015 July, 7]; Available from: [http://www.forever21.com/shop/ca/en/men-shirts-fitted/p/colorblocked-collar-shirt-2000057122--1001?utm\\_source=opmpros&utm\\_medium=affiliate&utm\\_campaign=UBAokkJVcos&siteID=UBAokkJVcos-UmsZe\\_R93G8VkdymPVM1YA](http://www.forever21.com/shop/ca/en/men-shirts-fitted/p/colorblocked-collar-shirt-2000057122--1001?utm_source=opmpros&utm_medium=affiliate&utm_campaign=UBAokkJVcos&siteID=UBAokkJVcos-UmsZe_R93G8VkdymPVM1YA).
33. ASOS. *Polo Ralph Lauren T-Shirt With Print*. [cited 2015 July, 7]; Available from: <http://www.asos.com/Polo-Ralph-Lauren/Polo-Ralph-Lauren-T-Shirt-With-Print/Prod/pgeproduct.aspx?iid=4361784&r=2>.
34. WeLoveShopping. *7th Street*. [cited 2015 July, 7]; Available from: <http://portal.weloveshopping.com/product/L90081559>.
35. J.CREW. *Ludlow Peak-Lapel Suit Jacket in Italina Worsted Wool*. [cited 2015 July, 7]; Available from: [https://www.jcrew.com/mens\\_category/suiting\\_m/ludlow/PRDOVR~06691/06691.jsp](https://www.jcrew.com/mens_category/suiting_m/ludlow/PRDOVR~06691/06691.jsp).
36. Forever21. *Slub Knit Sweater*. [cited 2015 July, 7]; Available from: [http://www.forever21.com/shop/ca/en/love21-sweaters-knits/p/slub-knit-sweater-2000056913--1001?utm\\_source=opmpros&utm\\_medium=affiliate&utm\\_campaign=UBAokkJVcos&siteID=UBAokkJVcos-002KHrjU5O\\_bOUBY8lz.sw](http://www.forever21.com/shop/ca/en/love21-sweaters-knits/p/slub-knit-sweater-2000056913--1001?utm_source=opmpros&utm_medium=affiliate&utm_campaign=UBAokkJVcos&siteID=UBAokkJVcos-002KHrjU5O_bOUBY8lz.sw).
37. TRITON. *Festee 2015 Triblend tank*. [cited 2015 July 7]; Available from: <http://to.ucsd.edu/collections/new-items/products/sgf15-festee-tank-triblend>.
38. informy. *Measurement Algorithms (Recall & Precision)*. 2012; Available from: <http://informy.tumblr.com/post/17650883277/measurement-algorithms-recall-precision>.





ภาคผนวก

จุฬาลงกรณ์มหาวิทยาลัย  
CHULALONGKORN UNIVERSITY

## ประวัติผู้เขียนวิทยานิพนธ์

นายวิศรุต สุรการินทร์ เกิดเมื่อวันที่ 29 พฤษภาคม พุทธศักราช 2533 ที่จังหวัดระยอง สำเร็จการศึกษาระดับปริญญาบัณฑิต หลักสูตรวิทยาศาสตรบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง เมื่อปีการศึกษา 2555 และเข้าศึกษาต่อในหลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมซอฟต์แวร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2556

