การจำแนกช่วงอายุบุคคลจากภาพถ่ายเต็มตัว

Age classification from full body images

นายชลิต เจริญศักดิ์
นายชวิศ ธัญญสัญชัย

โครงงานพิเศษนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต
สาขาวิชาวิทยาการคอมพิวเตอร์
คณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ปีการศึกษา 2556

การจำแนกช่วงอายุบุคคลจากภาพถ่ายเต็มตัว

Age classification from full body images

นายชลิต เจริญศักดิ์

นายชวิศ ธัญญสัญชัย

โครงงานพิเศษนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต

สาขาวิชาวิทยาการคอมพิวเตอร์

คณะวิทยาศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2556

Age classification from full body images

MR.CHALIT JAROENSAK

MR.CHAWIT THANYASANCHAI

A SPECIAL PROJECT SUBMITTED IN PERTIAL FULFILLMENT

OF THE REQUIREMENT OF THE DEGREE OF BACHELOR OF SCIENCE

IN COMPUTER SCIENCE

FACULTY OF SCIENCE

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

ACADEMIC YEAR 2013

หัวข้อโครงงานพิเศษ	การจำแนกช่วงอายุบุคคลจากภาพถ่ายเต็มตัว (Age classification from full body images)		
ชื่อนักศึกษา	นายชลิต	เจริญศักดิ์	53050965
	นายชวิศ	ธัญญสัญชัย	53050966
ปริญญา	วิทยาศาสตรบัณฑิต		
สาขาวิชา	วิทยาการคอมพิวเตอร์		
อาจารย์ที่ปรึกษา	อ.ธีระ	ศิริธีรากุล	

คณะวิทยาศาสตร์ สถาบัน เทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังอนุมัติให้โครงงานพิเศษนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร วิทยาศาสตรบัณฑิต สาขาวิทยาการคอมพิวเตอร์ ประจำปีการศึกษา 2556

คณะกรรมการสอบ	ลายมือชื่อ
ผศ.ดร.นวลสวาท หิรัญสกลวงศ์ (ประธานกรรมการ)	
รศ.ดร.จีรพร วีระพันธุ์ (กรรมการ)	
อ.ธีระ ศิริธีรากุล (กรรมการและอาจารย์ที่ปรึกษา)	

ลิขสิทธิ์ของคณะวิทยาศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

หัวข้อโครงงานพิเศษ การจำแนกช่วงอายุบุคคลจากภาพถ่ายเต็มตัว

ชื่อนักศึกษา นายชลิต เจริญศักดิ์ 53050965

นายชวิศ ธัญญสัญชัย 53050966

ปริญญา วิทยาศาสตรบัณฑิต

สาขาวิชา วิทยาการคอมพิวเตอร์

ปีการศึกษา 2556

อาจารย์ที่ปรึกษา **อ.ธีระ ศิริธีรากุล**

บทคัดย่อ

โฆษณานั้นถือว่าเป็นปัจจัยสำคัญอย่างหนึ่ง ที่ทำให้ผู้ประกอบการได้เปรียบทางด้านการค้ากับคู่แข่งทางการค้า แต่ในปัจจุบันโฆษณาตามสถานที่ต่างๆนั้นยังเป็นการโฆษณาโดยไม่มีการเลือกแสดงเฉพาะเจาะจงตามกลุ่มของลูกค้า ซึ่งในบางครั้งโฆษณาที่แสดงอาจไม่ตรงกับกลุ่มของลูกค้าที่อยู่บริเวณนั้น และเมื่อลูกค้าเห็นจึงไม่เกิดความสนใจ จึงทำให้เสียโอกาสทางการค้าไป ซึ่งเป็นจุดอ่อนของโฆษณาในปัจจุบันที่ไม่สามารถเลือกแสดงโฆษณาเฉพาะเจาะจงตามกลุ่มของลูกค้าได้ เราจึงได้พัฒนาโปรแกรมที่สามารถแบ่งกลุ่มลูกค้าตามช่วงอายุ มาช่วยให้โฆษณานั้นตรงกับกลุ่มลูกค้าเป้าหมายมากขึ้น เพื่อดึงดูดความสนใจของลูกค้าให้มากยิ่งขึ้น และช่วยในการตัดสินใจในการซื้อสินค้าและบริการของลูกค้าได้ดียิ่งขึ้น โดยจะเก็บข้อมูลภาพของบุคคลในช่วงอายุต่างๆกันสร้างเป็นData set และใช้กระบวนการเรียนรู้ของเครื่องเก็บคุณลักษณะต่างๆ แล้วจึงทำการตรวจจับและแบ่งช่วงอายุได้ด้วยเทคนิคที่นำHOG มาประยุกต์เพื่อทำการตรวจจับภาพบุคคล และ ใช้ SVM มาช่วยในการจำแนก

คำสำคัญ : OpenCV , image processing , Body Detection, Age and Gender Detection

Title	Age classification from full body images	
Student	Mr..Chalit Jaroensak	53050965
	Mr.Chawit Thanyasanchai	53050966
Degree	Bachelor of Science	
Major Program	**Computer Science**	
Academic Year	2013	
Advisor	**Mr.Teera Siriteerakul**	

ABSTRACT

Advertising in many locations is still not aim to the target customers that passed by. So, if we can specify range of age of customer and select the most appropriate advertising for that rage of age, then we will have an advantage over other competitions. Therefore, we collect data of people in various range of age for make Dataset and use the technique that developed from HOG to detect people and then use SVM to identify the target's range of age.

Keywords : OpenCV , image processing , Body Detection, Age and Gender Detection

กิตติกรรมประกาศ

ในการพัฒนาโครงงานปัญหาพิเศษการจำแนกช่วงอายุบุคคลจากภาพถ่ายเต็มตัว โดยอาศัยเทคนิคการตรวจจับร่างกาย คงไม่อาจสำเร็จได้เลย ถ้าหากไม่ได้รับความช่วยเหลือ และความร่วมมือจากหลาย ๆ ฝ่ายด้วยกัน บุคคลแรกที่ต้องกล่าวถึง ซึ่งเป็นส่วนสำคัญที่ทำให้โครงงานปัญหาพิเศษนี้เสร็จลงได้ ก็คือ อาจารย์ ธีระ ศิริธีรากุล ซึ่งเป็นอาจารย์ที่ปรึกษาปัญหาพิเศษ ที่ให้ความเอาใจใส่แนะนำแนวทางในการพัฒนา แนวทางในการแก้ปัญหา และช่วยเหลือเสมอมา ซึ่งต้องขอขอบพระคุณเป็นอย่างมาก รวมถึง อาจารย์ นวลสวาท หิรัญสกลวงศ์ และ อาจารย์ จีรพร วีระพันธุ์ ที่ช่วยตรวจสอบปัญหาพิเศษ รวมถึงอาจารย์ท่านอื่นๆที่ไม่ได้กล่าวถึง ณ ที่นี้ ที่ได้ให้การอบรมสั่งสอนตลอดเวลาที่ผ่านมา ขอขอบคุณ สาขาวิชาวิทยาการคอมพิวเตอร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ที่เอื้อเฟื้อสถานที่ และอุปกรณ์ต่าง ๆ สำหรับทำโครงงานปัญหาพิเศษ

ท้ายที่สุดต้องขอขอบพระคุณบุคคลสำคัญที่สุดที่ทำให้พวกข้าพเจ้ามีวันนี้ ก็คือ บิดา มารดา อันเป็นที่เคารพรักยิ่ง ซึ่งได้เลี้ยงดูพวกข้าพเจ้ามาเป็นอย่างดี พร้อมทั้งให้โอกาสในการศึกษาอย่างเต็มที่ และยังให้ กำลังใจ เอาใจใส่เสมอมา ในทุก ๆ ด้านอันหาที่เปรียบมิได้ ข้าพเจ้าขอระลึกในพระคุณอันสุดประมาณ และขอกราบขอบพระคุณมา ณ ที่นี้

นายชลิต เจริญศักดิ์

นายชวิศ ธัญญสัญชัย

สารบัญ

เรื่อง	หน้า
บทคัดย่อภาษาไทย	I
บทคัดย่อภาษาอังกฤษ	II
กิตติกรรมประกาศ	III
สารบัญ	IV
สารบัญตาราง	VII
สารบัญรูปภาพ	VIII
บทที่ 1 บทนำ	
1.1 ความเป็นมาและความสำคัญ	1
1.2 หลักการและเหตุผล	1
1.3 วัตถุประสงค์	2
1.4 ขอบเขตของงาน	2
1.5 ประโยชน์ที่คาดว่าจะได้รับ	2
1.6 ขั้นตอนในการดำเนินงาน	3
1.7 อุปกรณ์ที่ใช้ในการทำปัญหาพิเศษ	3
1.8 ขอบเขตและข้อจำกัดของโปรแกรมที่พัฒนา	4
1.8.1 ขอบเขตของโปรแกรมที่พัฒนา	4
1.8.2 ข้อจำกัดของโปรแกรมที่พัฒนา	4
บทที่ 2 ทฤษฎีที่เกี่ยวข้อง	
2.1 ฮิสโตแกรม (Histogram)	5
2.2 Histograms of Oriented Gradients for Human Detection (HOG)	6
2.3 Thresholding	9
2.4 Region Of Interest (ROI)	9
2.5 ระบบสี HSV (HSV Color)	10
2.6 LHSV (Local HSV)	11

เรื่อง	หน้า
2.7 Local Binary Pattern (LBP)	13
2.8 Local Direction Pattern (LDP)	14
2.9 Support Vector Machine (SVM)	16
2.10 เครื่องมือที่ใช้ในการพัฒนา	17
2.10.1 ภาษา C++	17
2.10.2 OpenCV (Open Source Computer Vision)	17
2.10.3 Data Set ของภาพบุคคลในช่วงอายุต่างๆ	18
2.10.4 กล้องWebcam	18
บทที่ 3 ขั้นตอนการพัฒนา	
3.1 รายละเอียดโปรแกรมที่จะพัฒนา (Software Specification)	19
3.1.1 Input/output Specification	19
3.1.2 Functional Specification	19
3.2 การเก็บข้อมูลเบื้องต้น	19
3.3 ขั้นตอนการวิจัย	23
3.3.1 ใช้ 1 คุณลักษณะ	23
3.3.2 ใช้ 2 คุณลักษณะ ร่วมกัน	24
3.4 โครงสร้างของซอฟแวร์ (Design)	25
บทที่ 4 ผลการดำเนินการ	
4.1 ผลการวิจัยหาวิธีการแยกแยะภาพบุคคล	27
4.1.1 การแยกแยะด้วย 1 คุณลักษณะ	27
4.1.2 การแยกแยะด้วย 2 คุณลักษณะ	28
4.2 ผลการพัฒนาโปรแกรม	30
บทที่ 5 สรุปผล และข้อเสนอแนะ	
5.1 สรุปผลการดำเนินงาน	36
5.2 ปัญหาในการดำเนินการ	37

เรื่อง	หน้า
5.3 ข้อเสนอแนะ	37
เอกสารอ้างอิง	38

สารบัญตาราง

ตารางที่	หน้า
2-1 การแบ่งช่วงทิศทางของ HOG	7
4-1 ผลการทดลองจากการใช้ 1 คุณลักษณะ	28
4-2 ผลการทดลองจากการใช้ 2 คุณลักษณะร่วมกัน	29

สารบัญรูปภาพ

รูปที่	หน้า
2-1 Histogram แสดงความสว่างของแสง	5
2-2 อธิบายลักษณะของ Histogram	6
2-3 ภาพตัวอย่างการแบ่ง blog และ cell ของ HOG	7
2-4 การเก็บค่า histogram ในแต่ละ blog โดยดูจากค่า gradient	8
2-5 ภาพรวมกระบวนการของ HOG	8
2-6 การแปลง Binary Image เมื่อกำหนด Threshold = 120	9
2-7 Region Of Interest (ROI)	10
2-8 Hue	10
2-9 Model HSV	11
2-10 Histogram HSV	12
2-11 การแบ่งช่วงค่า H ใน Histogram HSV	12
2-12 Local Binary Pattern	13
2-13 mask ของ Kirsch	14
2-14 กระบวนการ LDP	15
2-15 LDP Histogram	15
2-16 อัลกอริทึมของ Support Vector Machines	16
2-17 องค์ประกอบ OpenCV	18
3-1 ตัวอย่างภาพ Video Input	20
3-2 การ Detect บุคคลโดยใช้ HOG	20
3-3 การสร้าง Bounding Box	21
3-4 ภาพที่ได้จากการใช้ ROI	21
3-5 ตัวอย่าง Data Set ที่ได้	22
3-6 โครงสร้างของซอฟแวร์	25
4-1 หน้าหลักของโปรแกรม	30
4-2 กำลังตรวจจับช่วงอายุบุคคล	31
4-3 การแสดงของโฆษณา	31

4-4 กำลังตรวจจับช่วงอายุบุคล 32
4-5 ตรวจจับช่วงอายุบุคลสำเร็จ 32
4-6 การแสดงของโฆษณาของหน้าทดสอบการทำงานของโปรแกรม 32
4-7 หน้าการตั้งค่าหรือเปลี่ยนโฆษณา 33
4-8 หน้าการค้นหาและเลือกโฆษณา 33
4-9 หน้าแสดงเมื่อตั้งค่าหรือเปลี่ยนโฆษณา 33
4-10 หน้าแสดงข้อมูลของผู้พัฒนาและเทคนิคที่ใช้ในการทำ 34
4-11 การออกจากโปรแกรม 35

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญ

เนื่องจากในปัจจุบันมีการผลิตสินค้ามาแข่งขันกันมากมายในท้องตลาด ซึ่งในสินค้าและบริการ ชนิดหนึ่งนั้นได้มีการผลิตออกมาให้มีความเหมาะสมกับผู้บริโภคที่แตกต่างกันไป ได้แก่ เสื้อผ้า จะแบ่งออกเป็นเสื้อผ้าเด็ก เสื้อผ้าผู้ใหญ่ เป็นต้น

ดังนั้นการโฆษณา จึงเป็นสิ่งจำเป็นอย่างมากที่จะช่วยนำเสนอให้ผู้บริโภคทราบ และ เข้าใจถึงตัวสินค้าและบริการ เนื่องจากบางครั้งผู้บริโภคไม่ทราบถึงความต้องการของตนเอง โฆษณาจึงมีบทบาทสำคัญที่จะทำให้ผู้บริโภคเห็นความสำคัญของสินค้าและบริการ ซึ่งจะช่วยกระตุ้นและผลักดันให้เกิดการซื้อสินค้าและบริการ หรือ มีความต้องการใช้สินค้านั้นๆมากยิ่งขึ้นกว่าเดิม

จากปัญหาดังกล่าวจึงเป็นที่มาของการทำโครงงานปัญหาพิเศษนี้ขึ้นมา โดยนำทฤษฎีการประมวลผลภาพมาประยุกต์ เพื่อ วิเคราะห์ ช่วงอายุของมนุษย์โดยดูจากลักษณะร่างกายของมนุษย์ โดยการใช้ library ของ OpenCV และการเขียนโปรแกรมด้วยภาษา C++

1.2 หลักการและเหตุผล

ในปัจจุบันมีการโฆษณาโดยใช้สื่ออิเล็กทรอนิกส์ ตามบริเวณทางเดินในห้างสรรพสินค้าอย่างแพร่หลาย แต่ สำหรับสินค้าบางชนิด มีผลิตภัณฑ์ ที่เหมาะกับเพศและอายุ ที่ต่างกันไป ดังนั้นถ้าเราสามารถที่จะระบุได้ว่าลูกค้าที่เดินมา เป็นเด็ก ผู้ใหญ่ หรือ ผู้สูงอายุ ก็จะสามารถตอบสนองความต้องการและดึงดูด ผู้ที่เดินผ่านไป ผ่านมาได้มากยิ่งขึ้น ดังนั้นเราจึงพัฒนาโปรแกรม ซึ่งสามารถที่จะระบุคนที่เดินผ่านไปมาว่าเป็น เด็ก ผู้ใหญ่ผู้ชาย ผู้ใหญ่ผู้หญิง หรือ ผู้สูงอายุ เพื่อที่จะแสดงโฆษณาให้เหมาะกับกลุ่มลูกค้าให้ได้มากที่สุด

1.3 วัตถุประสงค์

- เพื่อศึกษาลักษณะความแตกต่างของช่วงอายุมนุษย์โดยดูจากรูปร่างของมนุษย์
- เพื่อศึกษาขั้นตอนและหลักการทำงานของการเรียนรู้ด้วยเครื่องคอมพิวเตอร์
- เพื่อทำการเก็บข้อมูลภาพมนุษย์แบบเต็มตัว เพื่อใช้ในการสร้าง Data Set
- เพื่อทำการวิเคราะห์และทดลองเพื่อตรวจสอบความแม่นยำและความถูกต้องในการตรวจจับช่วงอายุของมนุษย์

1.4 ขอบเขตของงาน

- ถ่ายวีดีโอคนที่เดินผ่านไปมาเพื่อที่จะนำไปสร้างเป็น Dataset ต่อไป
- Data Setที่ใช้จะต้องเป็นภาพของมนุษย์แบบเต็มตัว ซึ่งอยู่ในช่วงอายุต่างๆ
- ภาพที่ใช้ใน Data Set จะเป็นภาพมุมมองจากด้านหน้า ในระดับที่สูงกว่าสายตามนุษย์
- เมื่อตรวจจับได้จะแสดงผลเป็นโฆษณาตามที่กำหนดไว้สำหรับแต่ละช่วงอายุ
- ทำการทดลอง เพื่อหาเปอร์เซ็นต์ความถูกต้องในการตรวจจับช่วงอายุ แต่ละช่วงอายุ และสรุปผล
- ทำการสร้าง prototype ของโปรแกรม

1.5 ประโยชน์ที่คาดว่าจะได้รับ

สามารถนำผลการทดลองที่ได้ไปประยุกต์ใช้ เพื่อต่อยอด ไปใช้ในด้านการโฆษณาดังต่อไปนี้

1) นำเสนอโฆษณาได้เหมาะสมกับผู้บริโภคมากขึ้น
2) ดึงดูดความสนใจของผู้บริโภคให้เข้ามาชมโฆษณาของสินค้าและบริการ
3) ช่วยในการตัดสินใจในการซื้อของของผู้บริโภคได้ดียิ่งขึ้น

1.6 ขั้นตอนในการดำเนินงาน

1) วางแผนการดำเนินงานในการสร้างระบบตรวจจับรูปร่างและระบุช่วงอายุ
2) ศึกษาจากเอกสารและแหล่งข้อมูลที่เกี่ยวข้อง กับเนื้อหาในเรื่องเกี่ยวกับ การประมวลผลภาพ , OpenCV และ การเขียนโปรแกรมภาษา C++
3) เลือกวิธีการที่จะใช้ในการตรวจจับ
4) เก็บข้อมูลภาพนิ่งของมนุษย์แบบเต็มตัว
5) นำข้อมูลที่ได้มาสร้างเป็น Data Set ซึ่งแบ่งออกเป็นกลุ่มช่วงอายุต่างๆ ได้แก่ เด็ก ผู้ใหญ่ และ ผู้สูงอายุ
6) ดำเนินงานตามขั้นตอนที่วางแผนไว้ โดยใช้โปรแกรม Visual Studio 2010
7) ทดสอบการใช้งานระบบตรวจจับช่วงอายุ และประเมินผล
8) นำเสนอ และแสดงผลการทดลอง

1.7 อุปกรณ์ที่ใช้ในการทำปัญหาพิเศษ

1) เครื่องคอมพิวเตอร์
2) กล้อง Webcam
3) ภาพถ่ายเต็มตัวของมนุษย์ช่วงอายุต่างๆกัน จำนวนหนึ่ง
4) ซอฟต์แวร์ที่เกี่ยวข้อง เช่น OpenCV , Photoshop และ Visual Studio เป็นต้น

1.8 ขอบเขตและข้อจำกัดของโปรแกรมที่พัฒนา

1.8.1 ขอบเขตของโปรแกรมที่พัฒนา

- Data Setที่ใช้จะต้องเป็นภาพของมนุษย์แบบเต็มตัว ซึ่งอยู่ในช่วงอายุต่างๆ
- ภาพที่ใช้ในData Set จะเป็นภาพจริงของบุคคล มุมมองจากด้านหน้า ถ่ายในระดับที่สูงกว่าระดับสายตามนุษย์เล็กน้อย
- ทำการทดลอง เพื่อหาเปอร์เซ็นต์ความถูกต้องในการตรวจจับช่วงอายุ แต่ละช่วงอายุ และสรุปผล
- เมื่อตรวจจับได้จะแสดงผลเป็นโฆษณาตามที่กำหนดไว้สำหรับแต่ละช่วงอายุ

1.8.2 ข้อจำกัดของโปรแกรมที่พัฒนา

- ไม่สามารถตรวจจับช่วงอายุของคนที่เดินหันหลัง หรือเดินหันด้านข้าง
- ในการระบุช่วงอายุจะสังเกตจากเสื้อผ้าเครื่องแต่งกายเป็นหลัก ดังนั้นถ้ามีการใส่เสื้อผ้าคล้ายกัน อาจทำให้ระบุผิดพลาดได้
- ลักษณะของช่วงอายุผู้ใหญ่และผู้สูงอายุมีความใกล้เคียงกันมาก จึงอาจทำให้ระบุผิดพลาดได้

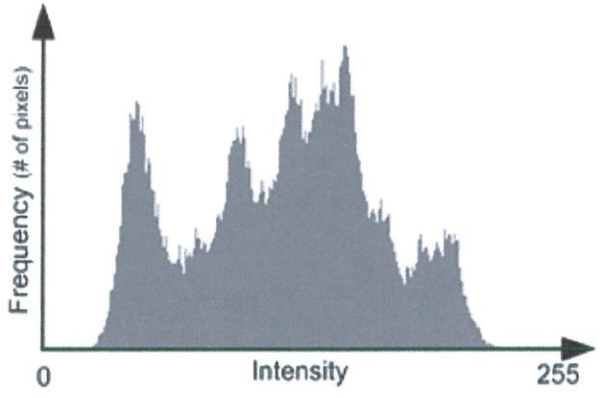
บทที่ 2

ทฤษฎีที่เกี่ยวข้อง

2.1 ฮิสโตแกรม (Histogram)

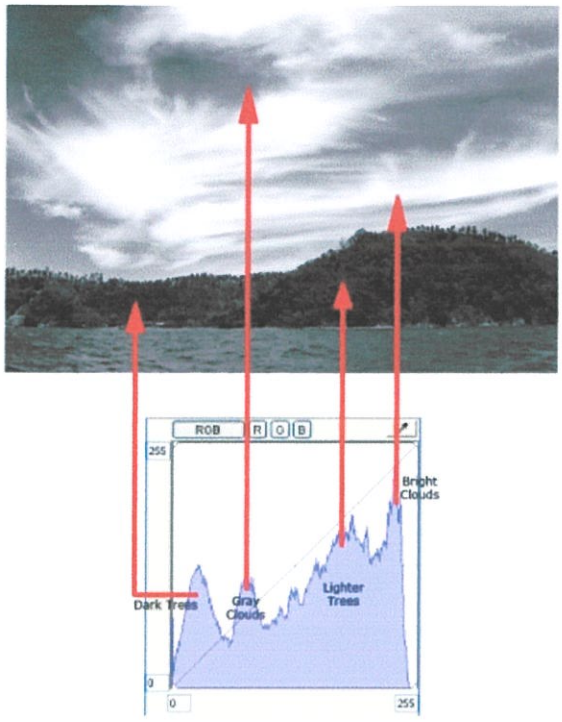
ฮิสโตแกรม เป็น กราฟที่ใช้ในการแสดงการกระจายของข้อมูล ซึ่งในด้านของImage Processing แล้ว กราฟฮิสโตแกรม จะใช้ในการสังเกตการณ์กระจายตัวของพิกเซลล์ของภาพ

สำหรับภาพ Gray scale กราฟฮิสโตแกรมจะแสดงถึง ระดับความสว่างของภาพ โดยในแนวแกน X จะระบุถึงความสว่างของภาพและในแนวแกน y จะระบุจำนวนพิกเซลล์ที่ตกที่ช่วงนั้นๆ ถ้าพิกเซลล์ตกที่ด้านซ้ายของกราฟมากแสดงว่าภาพนั้นค่อนข้างมืด และ ถ้าพิกเซลล์ตกทางขวาของกราฟมากแสดงว่าภาพมีความสว่างมาก



ภาพที่ 2-1 Histogram แสดงความสว่างของแสง

ที่มา: http://www.songho.ca/dsp/histogram/histogram.html



ภาพที่ 2-2 อธิบายลักษณะของ Histogram

ที่มา: http://www.howitookit.com/?p=185

2.2 Histograms of Oriented Gradients for Human Detection (HOG)

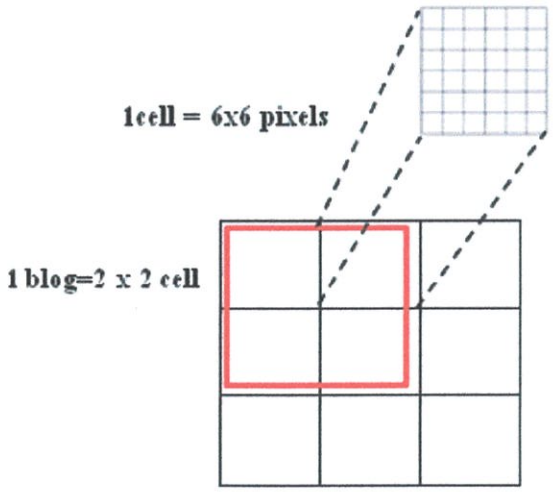
HOG เป็นวิธีการในการระบุคุณลักษณะของเส้นโดยใช้วิธีการเก็บข้อมูลHistogramของภาพ และนำไปเปรียบเทียบกับกลุ่มตัวอย่างเพื่อให้สามารถแยกแยะคุณสมบัติต่างๆได้ ซึ่งจะมีขั้นตอนคือ

1. แปลงภาพเป็น Gray Scale
2. คำนวณหาค่า Gradient ของภาพ โดยใช้วิธีของ Sobel ซึ่งใช้ mask [-1,0,1] และ $[-1,0,1]^T$ หรืออาจใช้ mask 3x3 ของ Sobel
3. หาค่า ขนาด และ ทิศทาง ของภาพ โดยจะทำการจัดแบ่งทิศทางเป็น 9 กลุ่ม ดังนี้

ช่วง / กลุ่ม	ช่วงแรก	ช่วงที่ 2
1	0-20 องศา	180-200 องศา
2	20-40 องศา	200-220 องศา
3	40-60 องศา	220-240 องศา
4	60-80 องศา	240-260 องศา
5	80-100 องศา	260-280 องศา
6	100-120 องศา	280-300 องศา
7	120-140 องศา	300-320 องศา
8	140-160 องศา	320-340 องศา
9	160-180 องศา	340-360 องศา

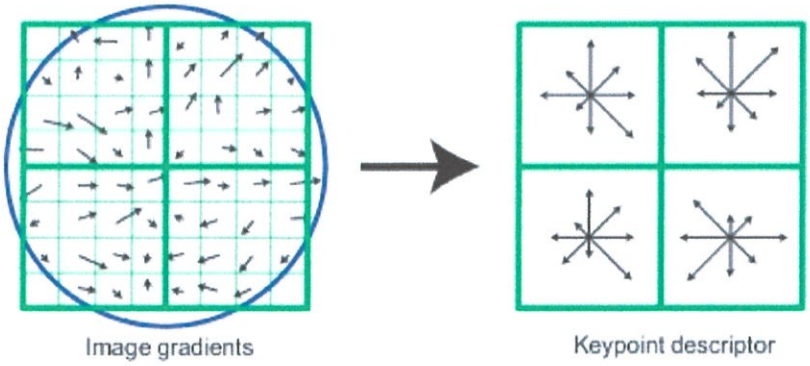
ตารางที่ 2-1 การแบ่งช่วงทิศทางของ HOG

4. แบ่งภาพโดยใช้ grid แบ่งภาพออกเป็น blog และ cell ขนาดเล็กซึ่งมีขนาดคงที่ (ปกติแล้วสำหรับมนุษย์จะใช้ขนาด cell ละ 6x6 pixel)



ภาพที่ 2-3 ภาพตัวอย่างการแบ่ง blog และ cell ของ HOG

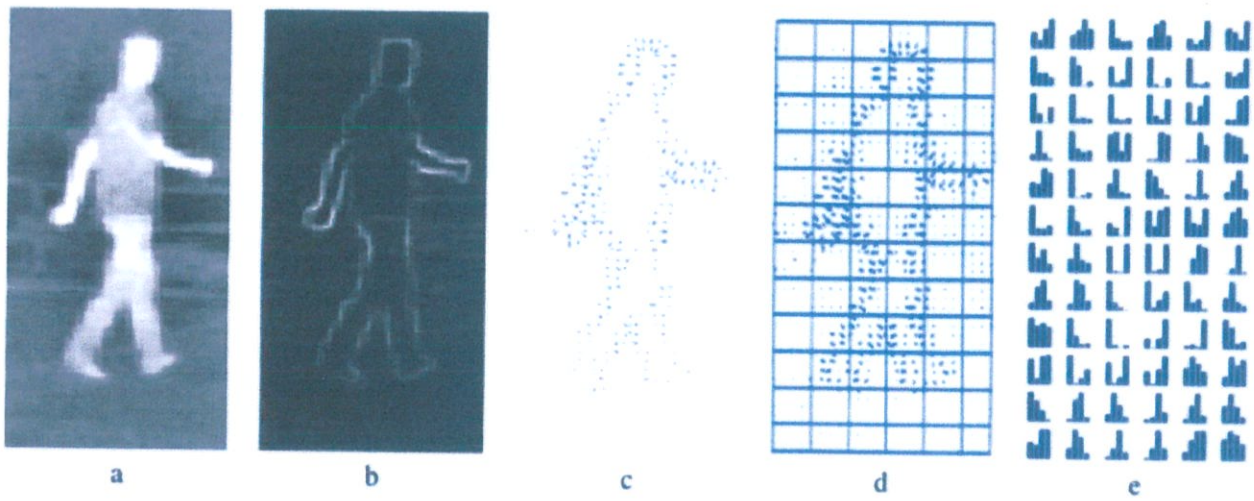
5. ทำการ Normalizeในส่วนของblogแต่ละblog เพื่อลดผลกระทบจากการเปลี่ยนแปลงของแสง และทำการเก็บค่าฮิสโตแกรมของแต่ละ blog



ภาพที่ 2-4 การเก็บค่า histogram ในแต่ละ blog โดยดูจากค่า gradient

ที่มา: http://homes.cs.washington.edu/~ankit/course_projects/files/vision/features/features.html

6. เข้าสู่กระบวนการของ Support Vector Machine ต่อไป

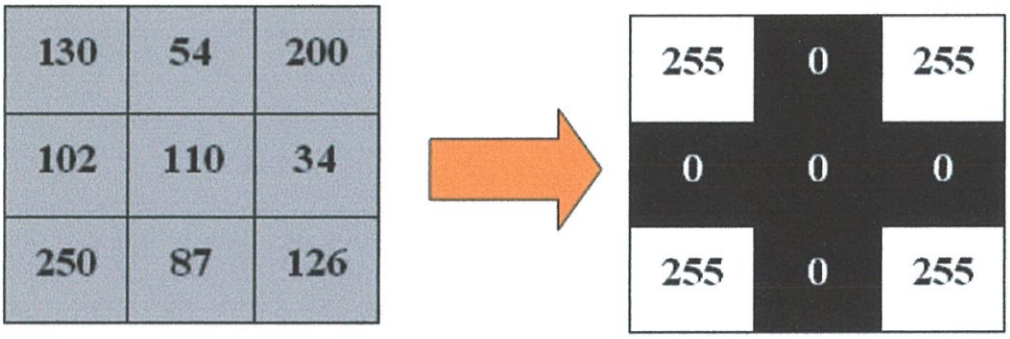


ภาพที่ 2-5 ภาพรวมกระบวนการของ HOG

2.3 Thresholding

Threshold (เทรสโช) เป็นค่าที่ถูกสมมุติขึ้น เพื่อใช้สำหรับเป็นเป็นจุดตรวจสอบ โดยค่า Threshold นี้มีส่วนสำคัญในการช่วยแปลงภาพเป็น Binary Image โดยการกำหนดค่าThreshodขึ้นมาค่าหนึ่ง แล้วนำไปเปรียบเทียบกับแต่ละpixel ถ้าpixelไหนมีค่ามากกว่าค่า Threshold ก็จะกำหนดให้ค่านั้นมีค่าเป็น255 หรือ สีขาว และถ้าpixelไหนมีค่าต่ำกว่าค่าThreshold ก็จะกำหนดให้ pixelนั้นมีค่าเป็น 0 หรือสีดำ

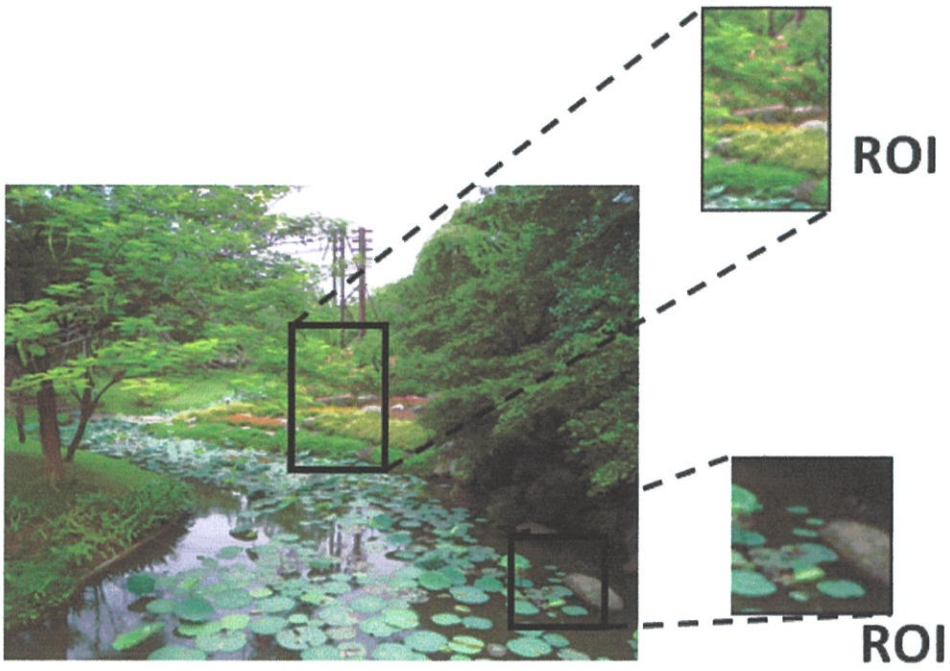
ตัวอย่างเช่นเมื่อกำหนด ค่า threshold ไว้ที่ 120 เพื่อจะทำการแปลงภาพ Gray Scale เป็นภาพขาวดำหรือ Binary Image



ภาพที่ 2-6 การแปลง Binary Image เมื่อกำหนด Threshold = 120

2.4 Region Of Interest (ROI)

Region Of Interest เป็นขั้นตอนในการหาบริเวณที่เราสนใจ โดย เมื่อระบุจุด coordinate (x , y) และความกว้างความยาวของบริเวณที่ต้องการได้แล้ว ก็จะทำการตีกรอบ ในบริเวณนั้น เพื่อต้องการที่จะนำภาพเฉพาะส่วนที่สนใจนั้นไปประมวลผลต่อไป ซึ่งในภาพ 1ภาพนั้น สามารถที่จะมี ROI กี่ภาพก็ได้

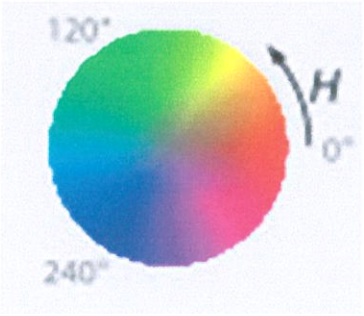


ภาพที่ 2-7 Region Of Interest (ROI)

2.5 ระบบสี HSV (HSV Color)

ระบบสี HSV ย่อมาจาก Hue, Saturation, Value ซึ่งระบบสี HSV จะใช้การพิจารณาจากองค์ประกอบ 3 อย่าง ได้แก่

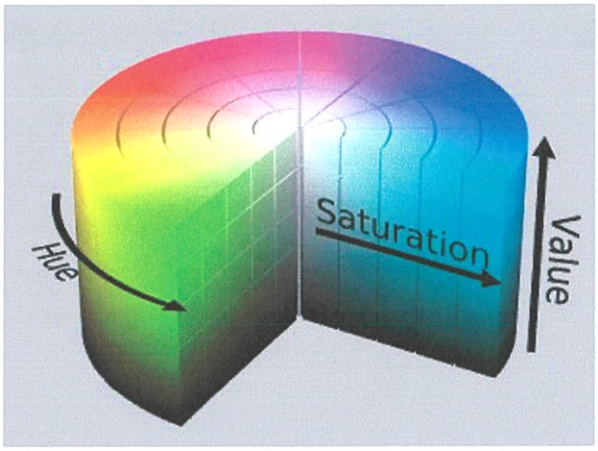
- Hue หรือ H คือองศาที่ห่างจากสีแดง โดยเริ่มจาก 0 องศา ถ้าค่า Hue = 0 จะแสดงว่าเป็นสีแดง



ภาพที่ 2-8 Hue

ที่มา: http://www.shadowwares.com/forum/index.php?topic=747.15

- Saturation หรือ S คือ ระยะห่างจากสีขาว (ระยะห่างจากแกนกลาง)
- Value หรือ V คือ ความสว่างของสีนั้นๆ



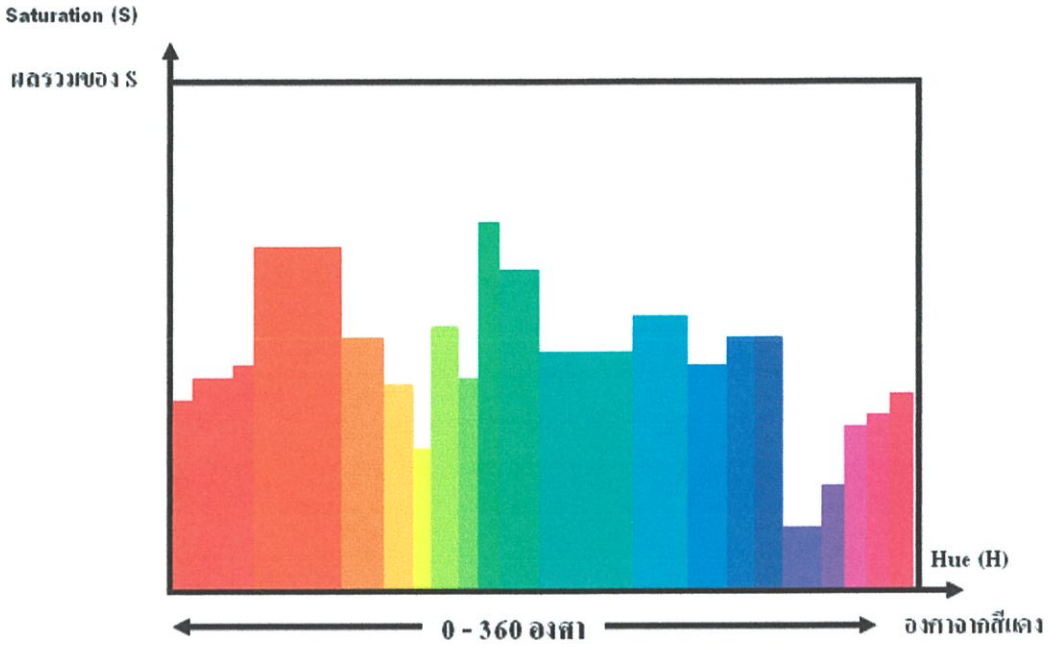
ภาพที่ 2-9 Model HSV

ที่มา: http://exceed.cpe.ku.ac.th/wiki/index.php/Color_Classify

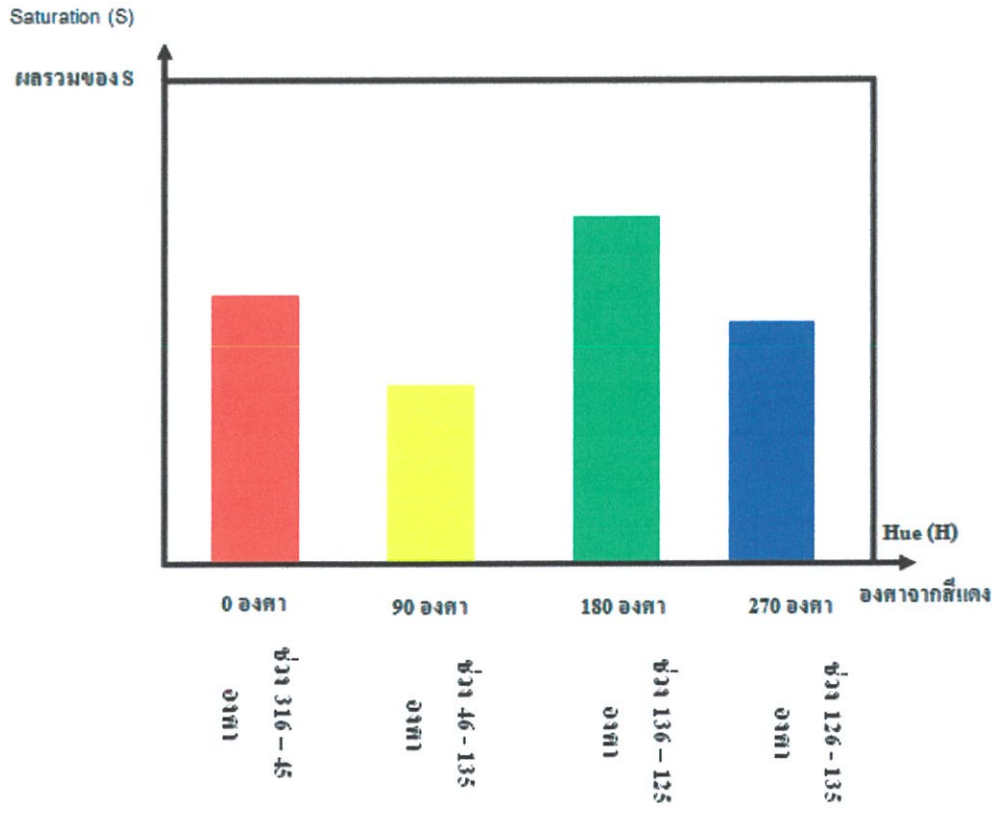
2.6 LHSV (Local HSV)

LHSV เป็นวิธีหนึ่งที่ใช้ในการแบ่งแยกคุณลักษณะของสีเป็นหลัก โดยวิธีนี้จะมีการทำงานคือ

1. แบ่งภาพออกเป็น blog ย่อยๆเช่นเดียวกับวิธี HOG
2. ทำการสร้าง Histogram ของระบบสี HSV จากแต่ละ blog โดยจะดูจากความสัมพันธ์ระหว่างค่า Hue(H) และค่า Saturation(S) โดยกำหนดให้แกน x ของกราฟ Histogram เป็นองศาของค่า Hue (H) และแกน y เป็น ค่าผลรวมของค่า Saturation(S) ดังนั้นจึงได้เป็นHistogramที่แสดงถึงผลรวมของค่า Saturation(S) ที่ตกในแต่ละช่วงของค่า Hue (H)
3. แบ่ง Histrogram ที่ได้ออกเป็น 4 bin
4. นำผลที่ได้จาก Histrogram ของทั้งภาพมาเรียงต่อกันเป็น vector เพื่อนำไปเข้าสู่กระบวนการเรียนรู้ของเครื่องต่อไป



ภาพที่ 2-10 Histogram HSV

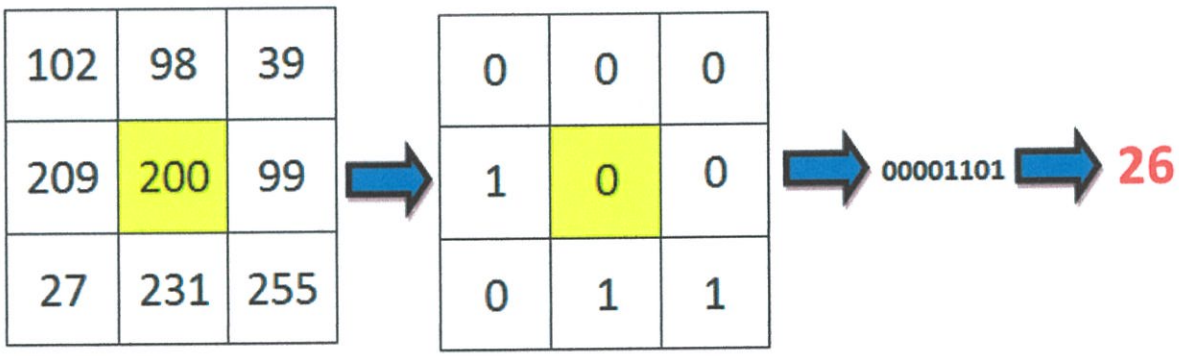


ภาพที่ 2-11 การแบ่งช่วงค่า H ใน Histogram HSV

2.7 Local Binary Pattern (LBP)

LBP เป็นวิธีในการเก็บคุณลักษณะของภาพแต่ละส่วน โดยจะใช้คุณสมบัติของจุดในภาพ โดยตัวดำเนินการของ LBP จะใช้ Matrix ขนาด 3x3 โดย

1. จะนำ ค่าของ pixel ตรงกลาง มาแปรียบเทียบกับ pixel รอบตัวมัน ถ้า pixel ไหนมีค่ามากกว่า pixel ตรงกลาง ก็จะกำหนดให้ pixel นั่นมีค่าเป็น 1 ส่วนถ้า pixel ไหน มีค่าน้อยกว่า pixel ตรงกลาง ก็จะกำหนดไหน pixel นั้นมีค่าเป็น 0
2. หลังจากนั้นกำหนดค่า pixel แล้ว จะนำค่าของแต่ละpixel โดยรอบ 8ค่า มาเรียงต่อกันเป็นเลขฐาน 2 โดยเริ่มจาก pixel บน-ซ้ายก่อน
3. ทำการแปลงเลขฐาน2 ที่ได้เป็นเลขฐาน10



ภาพที่ 2-12 Local Binary Pattern

4. หลังจากทำจนครบทั้งภาพแล้ว จะทำการแบ่งภาพเป็น blog ย่อยๆแล้วทำการเก็บค่า histrogram ของแต่ละ blog ซึ่งจะมีค่าที่เป็นไปได้ทั้งหมด 256 ค่า
5. นำค่าที่ได้ของทุก histrogram มาเรียงต่อกันเป็น vector เพื่อนำเข้าสู้การบวนการเรียนรู้ด้วยเครื่องต่อไป

2.8 Local Direction Pattern (LDP)

LDP เป็นการเก็บข้อมูลคุณลักษณะของภาพ โดย แบ่งภาพออกเป็นส่วนๆ และ ในแต่ละส่วนจะนำมากระทำกับ ตัวดำเนินการ ขนาด 3x3 ทั้งหมด 8 ตัว เพื่อเก็บคุณลักษณะของ 8 ทิศ

$$\begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix} \quad \begin{bmatrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{bmatrix} \quad \begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix} \quad \begin{bmatrix} 5 & 5 & -3 \\ 5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}$$

East M_0 North East M_1 North M_2 North West M_3

$$\begin{bmatrix} 5 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & -3 & -3 \end{bmatrix} \quad \begin{bmatrix} -3 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix} \quad \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix} \quad \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{bmatrix}$$

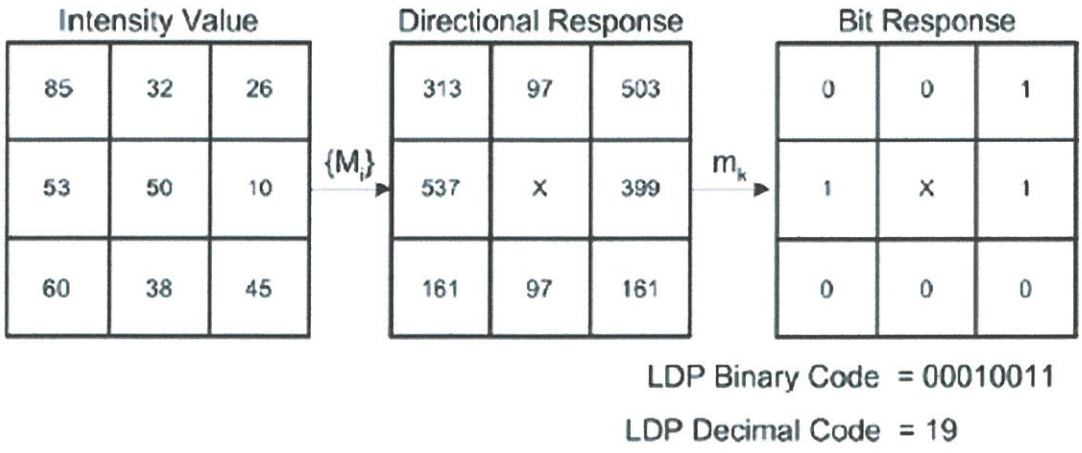
West M_4 South West M_5 South M_6 South East M_7

ภาพที่ 2-13 mask ของ Kirsch

ที่มา : Hasanul Kabir, Taskeed Jabid, and Oksam Chae.

โดย LDP มีกระบวนการทำงานดังนี้

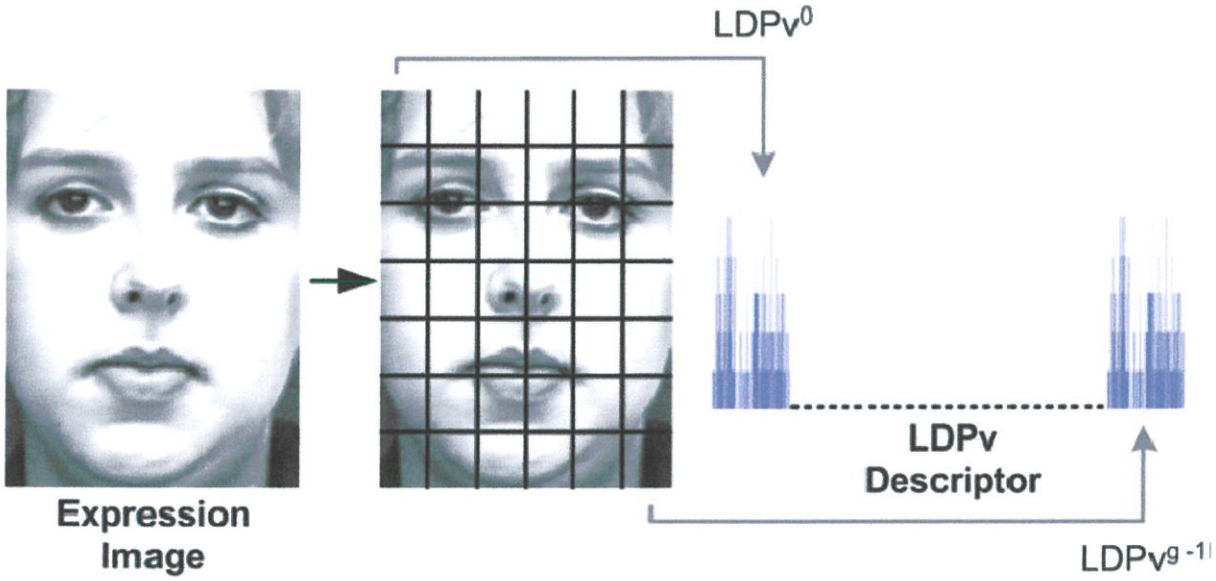
1. นำ mask ทั้ง 8 ตัว มาทำการ convolution กับภาพ
2. นำผลลัพธ์ที่ได้จากทั้ง 8 ทิศ มาทำการเปรียบเทียบกัน แล้วกำหนดให้ ค่าที่มากที่สุด 3 ตัวมีค่าเป็น 1 และค่าอื่น มีค่าเป็น 0
3. จากนั้นนำค่าที่ได้มาเรียงต่อกันเป็นเลขฐาน 2



ภาพที่ 2-14 กระบวนการ LDP

ที่มา : Hasanul Kabir, Taskeed Jabid, and Oksam Chae.

4. หลังจากทำจนครบทั้งภาพแล้ว จากนั้นจะทำการแบ่งภาพเป็น blog ย่อยๆ แล้วจึงทำการเก็บค่า histrogram ของแต่ละ blog แล้วนำค่า histrogram ทั้งหมดมาเรียงต่อกัน เป็น vector เพื่อเข้าสู่กระบวนการเรียนรู้ด้วยเครื่องต่อไป

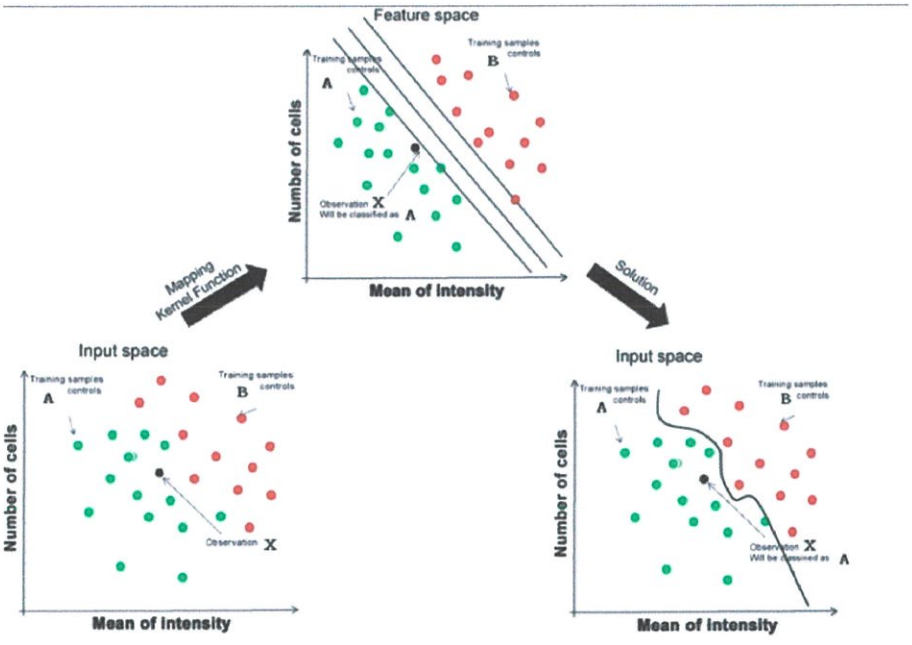


ภาพที่ 2-15 LDP Histogram

ที่มา : Hasanul Kabir, Taskeed Jabid, and Oksam Chae.

2.9 Support Vector Machine (SVM)

เป็นเทคนิคการเรียนรู้ของเครื่อง (Machine Learning) ในรูปแบบการเรียนรู้เชิงสถิติ ซึ่งจะเป็นการเรียนรู้แบบมีผู้สอน (Supervised learning) ซึ่งจะนำมาช่วยในการแก้ปัญหาการจัดประเภทข้อมูล (Classification) ที่มีความสลับซับซ้อนได้อย่างมีประสิทธิภาพ โดยจะใช้ระนาบแบ่งเขตข้อมูลที่เหมาะสม (The Optimal Separating Hyperplane)
เป็นตัวจัดประเภทข้อมูล ซึ่งจะมีโครงสร้างแบบเส้นตรง และสามารถสร้างพื้นที่ระยะห่างระหว่างตัวจัดประเภทข้อมูลเองกับค่าที่ใกล้ที่สุดของแต่ละกลุ่มข้อมูลได้มากที่สุด เพื่อประสิทธิภาพในการแยกประเภทของชุดข้อมูลแต่ละประเภทออกจากกันอย่างชัดเจน โดยหลักการทำงานเพื่อจำแนกประเภทข้อมูลของ Support Vector Machine จะได้ดังภาพ



ภาพที่ 2-16 อัลกอริทึมของ Support Vector Machines

ที่มา :

https://www.eleceng.adelaide.edu.au/personal/dabbott/wiki/index.php/Final_Report_2011:_Who_wrote_the_Letter_to_the_Hebrews%3F

2.10 เครื่องมือที่ใช้ในการพัฒนา

2.10.1 ภาษา C++

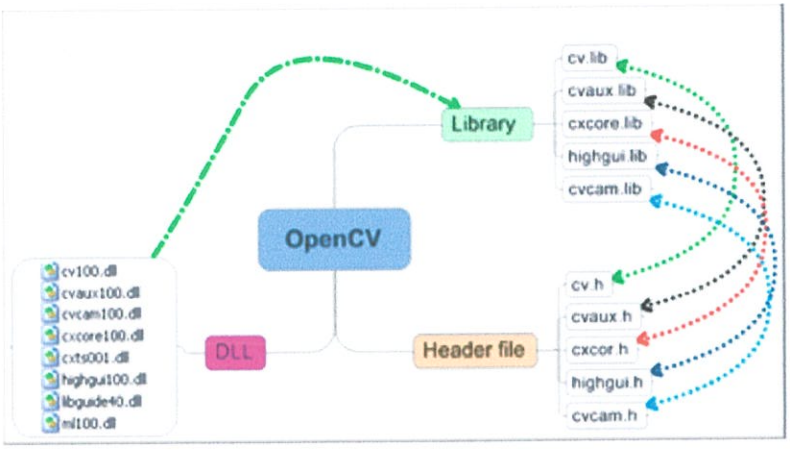
เป็นภาษาที่ใช้ในการเขียนโปรแกรมทั่วไป รองรับการเขียนโปรแกรมหลากหลายรูปแบบ (multi-paradigm) ซึ่งภาษา C++ ถูกออกแบบมาให้ทำงานร่วมกับภาษา C เกือบทุกกรณี และ C++ ยังออกแบบมาให้สามารถทำงานบน platform ที่ต่างกันได้ด้วย (multi-platform)

ในการพัฒนาโปรแกรมนี้ เราจึงทำการเขียนด้วยภาษา C++ บนโปรแกรม Visual Studio 2010

2.10.2 OpenCV (Open Source Computer Vision)

OpenCV เป็น Open Source ซึ่งใช้ในด้านการประมวลผลภาพ โดยมีจุดเด่นที่สามารถเรียกใช้ library ที่ OpenCV เตรียมไว้ให้ในการประมวลผลภาพได้ ทั้งภาพนิ่งและภาพเคลื่อนไหว ซึ่งตัว library ของ OpenCV นี้สามารถเรียกใช้ได้ ทั้งในภาษา C, C++, C# และ JAVA เป็นต้น จึงทำให้ OpenCV ใช้งานได้สะดวก และสามารถพัฒนาได้ด้วยภาษาโปรแกรมที่หลากหลาย ซึ่ง OpenCV นี้มีองค์ประกอบสำคัญ 4 ส่วน ดังนี้

1. **CXCORE** เป็นฟังก์ชันเบื้องต้นที่ใช้จัดการเกี่ยวกับจุด ขนาด อาร์เรย์ หน่วยความจำคำสั่งในการวาดภาพ และการประกาศตัวแปรภาพ
2. **CV** ใช้ในการประมวลผลและการวิเคราะห์รูปภาพ
3. **Machine Learning** เป็นไลบรารีที่รวมคลาสและฟังก์ชันทางสถิติ (Statistical) การแยกคลาสและการแบ่งกลุ่มของข้อมูล (Clustering)
4. **HighGUI** เป็นไลบรารีที่ใช้ในการดึงภาพ การบันทึกภาพ การเปลี่ยนขนาดและเคลื่อนย้ายหน้าต่าง รวมไปถึงการ ตรวจสอบเมาส์ และแป้นพิมพ์



ภาพที่ 2-17 องค์ประกอบ OpenCV

ที่มา: http://kwangee1245.blogspot.com/

2.10.3 Data Set ของภาพบุคคลในช่วงอายุต่างๆ

Data Set ที่จะใช้สำหรับโครงการนี้ จะได้จากการเก็บข้อมูล โดยใช้กล้องถ่าย Video และนำมาจับภาพและคัดแยกออกเป็นช่วงอายุต่างๆ

2.10.4 กล้อง Webcam

กล้องที่ใช้ในการรับภาพเข้ามาประมวลผลสำหรับโปรแกรมนี้ จะใช้กล้อง Webcam ซึ่งมีความละเอียดไม่สูงมากนักเพื่อให้เหมาะสมกับ Data Set ที่ใช้

บทที่ 3

ขั้นตอนการพัฒนา

3.1 รายละเอียดโปรแกรมที่จะพัฒนา (Software Specification)

3.1.1 Input/output Specification

Input : ภาพ video ของบุคคลที่เดินผ่านไปมา

Output : แสดงเป็นโฆษณาซึ่งปรับเปลี่ยนตามช่วงอายุของบุคคลที่จับภาพได้

3.1.2 Functional Specification

- ทำการจับภาพของบุคคลที่เดินผ่านไปมาได้แบบ Real time
- นำภาพที่จับได้มาทำการวิเคราะห์และแยกแยะช่วงอายุของบุคคลว่าอยู่ในช่วงอายุใด ได้แก่ เด็ก, ผู้ใหญ่ผู้ชาย , ผู้ใหญ่ผู้หญิง และ ผู้สูงอายุ
- นำผลที่ได้ไปปรับการแสดงผล ให้แสดงผลเป็นไฟล์ video โฆษณา ตามที่ได้กำหนดไว้

3.2 การเก็บข้อมูลเบื้องต้น

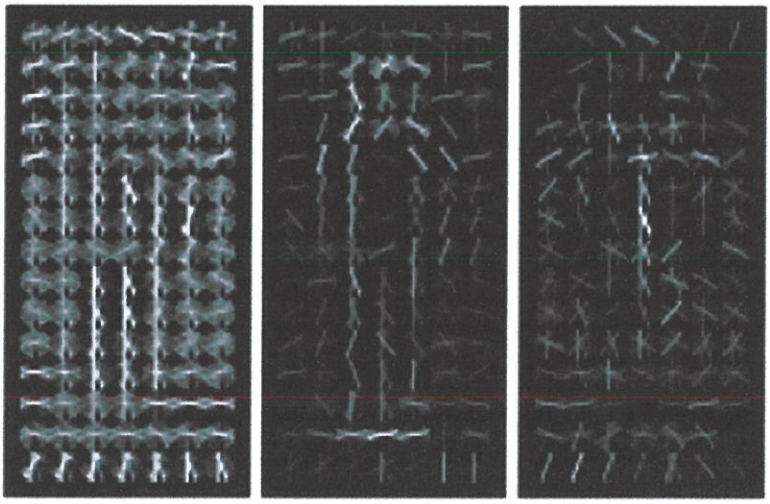
ในการเก็บข้อมูล จะทำโดยการนำกล้องวีดีโอไปตั้งถ่ายไว้ที่บริเวณใดบริเวณหนึ่ง โดยที่ข้อมูลในDataset ควรเก็บข้อมูลมาจากมุมมอง และสถานที่เดียวกันทั้งหมด เนื่องจากจะช่วยลดปัญหาสภาพแวดล้อมที่แตกต่างกันของข้อมูล ซึ่งจะส่งผลให้ ความแม่นยำเมื่อนำไปใช้ ตกลงด้วย ซึ่งโดยทั่วไปแล้วสถานที่ที่เหมาะสมแก่การเก็บข้อมูลมากที่สุดสถานที่หนึ่ง ได้แก่ บนสะพานลอย เนื่องจาก บนสะพานลอย เราสามารถที่จะกำหนดเส้นทางเดินของคนได้ง่าย รวมทั้งสามารถที่จะ เก็บภาพในมุมมองที่เราต้องการได้ง่ายด้วย

กระบวนการที่จะทำในขั้นตอนการเก็บข้อมูลมีดังต่อไปนี้

(1) รับภาพ Input ซึ่งเป็นภาพ Video เข้ามา

ภาพที่ 3-1 ตัวอย่างภาพ Video Input

(2) ใช้ library HOG Default Human Detection ของ OpenCV ตรวจจับว่า บริเวณใดของภาพคือภาพของบุคคล



ภาพที่ 3-2 การ Detect บุคคลโดยใช้ HOG

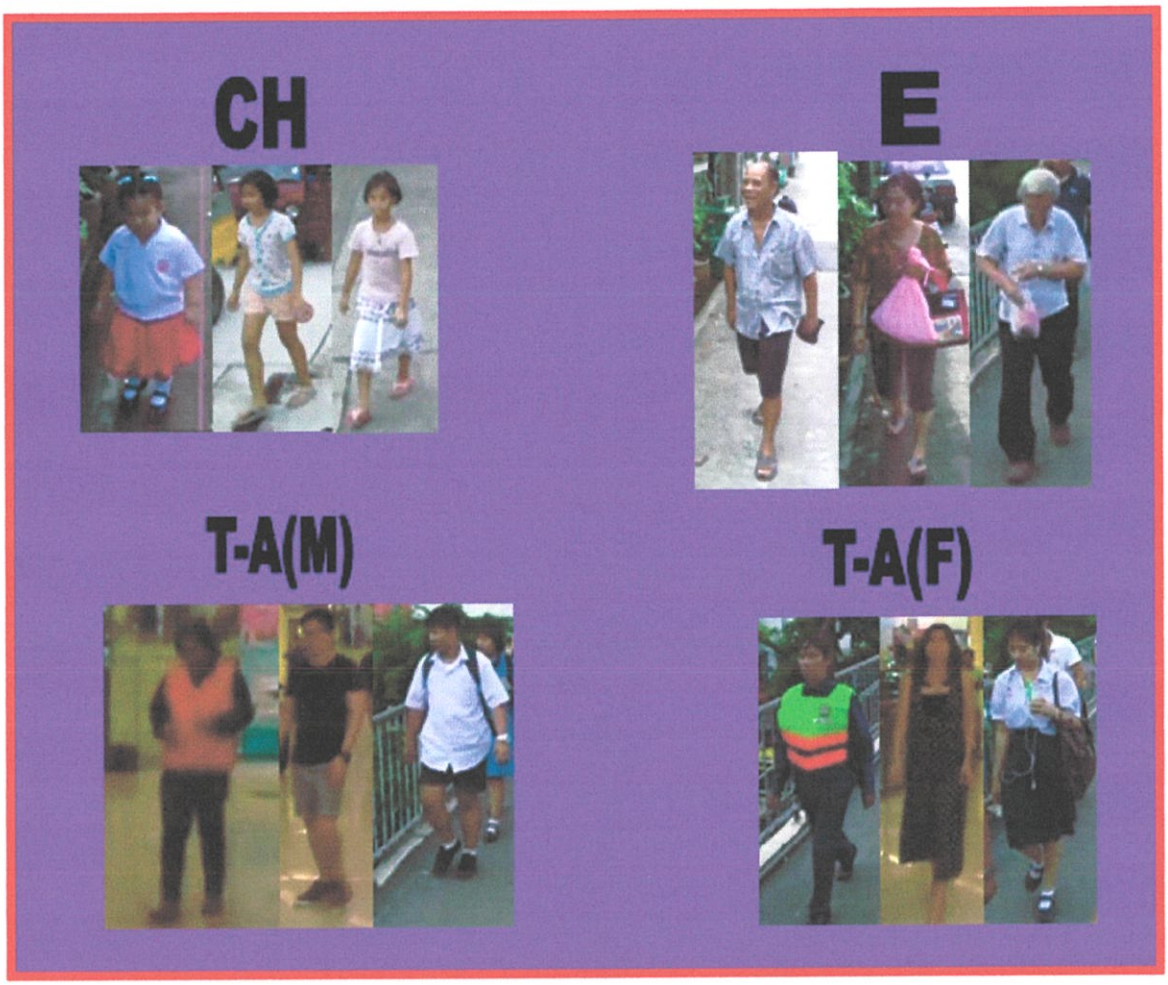
ที่มา: http://cs.brown.edu/courses/cs143/2011/proj4/

(3) เมื่อตรวจจับได้แล้วว่าบริเวณใดคือภาพบุคคล จะทำการเก็บค่า coordinate และขนาด ของบริเวณนั้นเอาไว้ เพื่อทำการสร้าง bounding box มาครอบไว้

ภาพที่ 3-3 การสร้าง Bounding Box

(4) จากนั้นใช้ ROI เพื่อตัดภาพบริเวณที่สนใจออกมา จากนั้นจึงทำการ Save ภาพเก็บไว้ เพื่อนำไปคัดแยกเพื่อทำ Data Set ต่อไป

ภาพที่ 3-4 ภาพที่ได้จากการใช้ ROI



ภาพที่ 3-5 ตัวอย่าง Data Set ที่ได้

ในการคัดแยก Dataset จะทำการคัดแยกด้วยสายตา โดยแบ่งออกเป็น 4 กลุ่ม ได้แก่ เด็ก , วัยรุ่น-ผู้ใหญ่ ผู้ชาย, วัยรุ่น-ผู้ใหญ่ ผู้หญิง และ ผู้สูงอายุ ซึ่งการแยกด้วยสายตา จะไม่สามารถแยกช่วงอายุได้อย่างละเอียด จึงจำเป็นต้องแยกช่วงอายุที่สามารถเห็นได้ชัดเจน เท่านั้น คือ ช่วงเด็ก จะเป็นกลุ่มข้อมูลที่เป็นกลุ่มเด็กเล็ก คือ อายุประมาณไม่เกิน 14 ปี และช่วงของ ผู้สูงอายุ จะแยกจากภาพที่เห็นได้ชัดที่สามารถระบุได้ว่าเป็นผู้สูงอายุเท่านั้น

จากนั้นจึงนำ Dataset ที่ได้มา เข้าสู่กระบวนการTraining ของ SVM (Support Vector Machine) เพื่อสอนเครื่อง (machine learning) ว่าภาพลักษณะใดควรจัด อยู่ในกลุ่มของช่วงอายุใด

3.3 ขั้นตอนการวิจัย

เราต้องทำการทดลองเพื่อที่จะหาวิธีการ ซึ่งให้ผลความแม่นยำ สูงที่สุดเพื่อที่จะนำวิธีการดังกล่าวไปใช้จริงในการพัฒนาโปรแกรมต่อไป

โดยวิธีการที่จะนำมาใช้ในการทดลอง มี 4 วิธีด้วยกัน ได้แก่

- Histograms of Oriented Gradients for Human Detection (HOG)
  ใช้ในการเก็บคุณลักษณะของเส้น
- Local Binary Pattern (LBP)
  ใช้ในการเก็บจุดเด่นของภาพ
- Local Direction Pattern (LDP)
  คล้ายกับวิธี LBP แต่จะแบ่งการเก็บคุณลักษณะจุดเด่นของภาพออกเป็น 8 ทิศ
- Local HSV (LHSV)
  ใช้เก็บคุณลักษณะของสี ในระบบสี HSV

ในการทดลองนี้จะทำการทดลองด้วยวิธี n-fold โดยจะทำการแบ่งกลุ่มของข้อมูลออกเป็น 10ชุด แล้วทำการทดสอบ 10 รอบ โดยในแต่ละรอบจะแบ่งชุดข้อมูลออกเป็นชุดข้อมูลที่ใช้ในการทดสอบความถูกต้อง และชุดข้อมูลที่จะใช้ในการเรียนรู้ของเครื่อง ซึ่งรายละเอียดในส่วนนี้จะได้กล่าวถึงในบทถัดไป

การทดลองจะแบ่งเป็น 2 ช่วง ได้แก่ การทดลองโดยใช้วิธีการใดวิธีการหนึ่งเพียงวิธีเดียว และ การนำ 2 วิธีมาใช้ร่วมกัน

3.3.1 ใช้ 1 คุณลักษณะ

ขั้นตอนการ training

ทำโดยนำภาพใน Data set ในชุดที่ใช้ในการเรียนรู้ของเครื่อง มาแปลงเป็น vector ด้วยวิธีการใดวิธีการหนึ่งใน 4 วิธี และนำ vector ที่ได้ เข้าสู่กระบวนการของ SVM เพื่อที่จะสอนเครื่อง โดยติด tag ไว้ที่แต่ละ vector เพื่อสอนเครื่องว่า vector ไหนควรจัดอยู่ในช่วงอายุใด

ขั้นตอนการ Predict

ในขั้นตอนนี้จะใช้ภาพของบุคคลในชุดทดสอบความถูกต้อง โดยเราจะทำการกำหนดค่าไว้ก่อนว่าภาพนั้นๆเป็นภาพของบุคคลเพศใด หรือ ในช่วงอายุใด จากนั้นจึงนำภาพนั้นมาแปลงเป็น vector ด้วยวิธีการใดวิธีการเดียวกันกับวิธีที่ใช้ในขั้นตอน training จากนั้นจึงนำ vector ที่ได้ เข้าสู่กระบวนการของ SVM โดย ระบบจะดูว่า vector ที่ได้ตกอยู่ในกลุ่มข้อมูลกลุ่มใดใน SVM แล้วจึงนำผลที่ได้มาเปรียบเทียบกับค่าที่เราได้กำหนดไว้ในตอนแรก ว่าถูกต้องหรือไม่ โดยจะทำเช่นเดียวกันนี้ กับ ภาพทุกภาพในชุดทดสอบ แล้วจึงสรุปหาเปอร์เซ็นต์ความถูกต้องต่อไป

3.3.2 ใช้ 2 คุณลักษณะ ร่วมกัน

การทดลองด้วย 2 features จะทำโดยการใช้ วิธี HOG , LBP , LDP ซึ่งเป็นวิธีที่ใช้ในการเก็บคุณลักษณะของเส้น หรือจุดเด่นของภาพ มาใช้ร่วมกับ วิธี LHSV ซึ่งเก็บคุณลักษณะของสี เพื่อให้สามารถเก็บได้ทั้ง คุณลักษณะด้านรูปร่าง และคุณลักษณะของสีด้วย จึงแยกการทดลองได้เป็น 3 วิธี ได้แก่

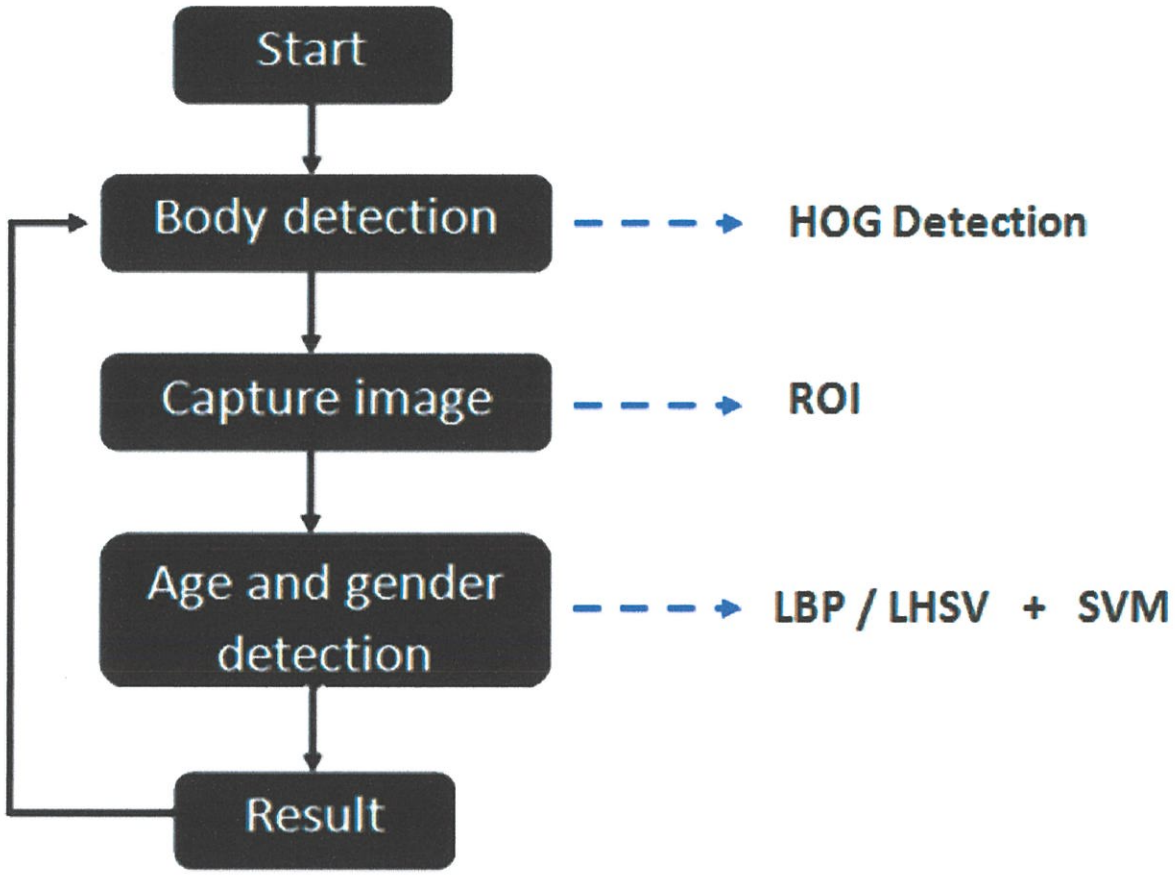
- HOG + LHSV
- LBP + LHSV
- LDP + LHSV

ซึ่งสำหรับการทดลองโดยใช้ 2 คุณลักษณะ ร่วมกันนี้ ขั้นตอนวิธีการทำจะเหมือนกับการทดลองด้วย 1 คุณลักษณะ แต่ในขั้นตอนการแปลงภาพเป็น vector จะทำโดยการหา vector ด้วยวิธีแรกก่อน จากนั้นหาvector ของภาพด้วยวิธี LHSV จากนั้น จึงนำ vector ที่ได้ทั้ง2vector มาต่อกัน เป็น vector เดียว แล้วจึงนำ vector ที่ได้นั้นเข้าสู่กระบวนการของ SVM

จากผลการทดลองที่ได้จะสามารถ สรุปได้ว่า วิธีการที่ให้ความแม่นยำสูงที่สุด ได้แก่วิธี LBP + LHSV ซึ่งมีความแม่นยำ อยู่ที่ 66.8 % รายละเอียดในส่วนนี้จะได้กล่าวถึงในบทถัดไป

เนื่องจาก กลุ่มของ Dataset ถูกแบ่งเป็น 4 กลุ่ม ความแม่นยำจากการเดาสุ่มปกติจึงอยู่ที่ 25% จากผลที่ได้ความแม่นยำ 66.8 % ถือว่าเป็นความแม่นยำที่น่าพึงพอใจในระดับหนึ่ง

3.4 โครงสร้างของซอฟแวร์ (Design)



ภาพที่ 3-6 โครงสร้างของซอฟแวร์

โปรแกรมจะทำงานโดยเริ่มจาก

- เมื่อได้รับภาพจากวีดีโอ เข้ามาแล้ว จะทำการใช้ library ของ OpenCV ซึ่งได้เตรียมฟังก์ชัน HOG Default Human Detection ไว้ให้อยู่แล้ว ทำให้สามารถเรียกใช้ เพื่อตรวจจับได้ทันทีว่าบริเวณใดของภาพคือภาพมนุษย์

- เมื่อตรวจจับได้แล้วว่า บริเวณใด คือภาพของมนุษย์ ก็จะทำการเก็บจุด coordinate ไว้ แล้วจึงใช้ ROI ตัดภาพบริเวณนั้นออกมา
- ภาพของบุคคลที่ถูกตัดออกมา จะถูกนำมาแปลงเป็น vector ด้วยวิธีการใด วิธีการหนึ่ง ได้แก่ HOG , LBP , LDP หรือ LHSV ซึ่งในที่นี้ได้เลือกใช้วิธี LBP ร่วมกับ LHSV เนื่องจากจากผลการทดลองที่ได้วิธีการ LBP + LHSV เป็นวิธีที่มีความแม่นยำสูงที่สุด โดยการใช้ 2 วิธี ร่วมกันนี้ จะทำโดยการ หา vector ของภาพด้วยวิธี LBP ก่อน แล้วจากนั้นจึงหา vector ของภาพด้วยวิธี LHSV แล้วนำ vector ที่ได้ จากทั้ง 2 วิธี มาต่อกันเป็นvector เดียว
- จากนั้นจึงนำ vector ที่ได้ เข้าสู่กระบวน การของ SVM เพื่อ ทำการ Predict ผลต่อไปว่าภาพที่ได้รับมานั้น เป็นภาพของบุคคล เพศใด หรือ ช่วงอายุใด
- เมื่อได้ผลออกมาแล้วก็จะทำการแสงผลออกมาเป็นโฆษณา ตามที่ได้ กำหนดไว้ตามช่วงอายุต่างๆ
- เมื่อโฆษณาถูกแสดงจนจบแล้วระบบ ก็จะวนกลับมาเข้าสู่กระบวนการในการ ตรวจจับ ภาพบุคคล เพื่อที่จะนำไปประมวลผลในรอบต่อไป

บทที่ 4

ผลการดำเนินการ

4.1 ผลการวิจัยหาวิธีการแยกแยะภาพบุคคล

4.1.1 การแยกแยะด้วย 1 คุณลักษณะ

ในการทดสอบความถูกต้องจากการเรียนรู้ด้วยเครื่องและการรู้จำ ได้ใช้วิธีต่างๆดังนี้

- HOG
- LBP
- LDP
- LHSV

โดยในการทดลองจะนำผลความถูกต้องจากการทดลองแต่ละวิธีมาเปรียบเทียบกัน โดยชุดของ ข้อมูลที่ใช้ในการทดลองที่ใช้จะเป็นชุดข้อมูลที่ได้จากการเก็บข้อมูลภาพจากคนที่เดินบนสะพานลอยในเวลากลางวัน จำนวน 2180 ภาพ ซึ่งในการทดลองจะทำการแบ่งกลุ่มของข้อมูลออกเป็น 4 กลุ่ม คือกลุ่มของเด็ก, กลุ่มของวัยรุ่น-ผู้ใหญ่ ผู้ชาย, กลุ่มของวัยรุ่น-ผู้ใหญ่ ผู้หญิง และกลุ่มของผู้สูงอายุ

ซึ่งในการทดลองนี้จะทำการทดลองด้วยวิธี n-fold โดยจะทำการแบ่งกลุ่มของข้อมูลออกเป็น 10ชุด แล้วทำการทดสอบ 10 รอบ โดยแต่ละรอบจะแบ่งออกเป็น ชุดที่ใช้ทดสอบ และใช้ชุดที่ใช้ในการฝึกฝน ซึ่งแบ่งเป็น

- ชุดทดสอบ 217 ภาพ เป็น
  เด็ก 10 ภาพ ผู้ชาย 70 ภาพ ผู้หญิง 124 ภาพ ผู้สูงอายุ 13 ภาพ
- ชุดที่ใช้ในการฝึกฝน 1963 ภาพ เป็น
  เด็ก 97 ภาพ ผู้ชาย 631 ภาพ ผู้หญิง 1117ภาพ ผู้สูงอายุ 118 ภาพ

ซึ่งจากการทดลองได้ผลการทดลองดังนี้

วิธีที่ใช้	%เฉลี่ยความแม่นยำจากการทดลอง	ส่วนเบี่ยงแบนมาตรฐาน
HOG	59.6	3.3
LBP	56.2	7.4
LDP	59.7	4.1
LHSV	43.8	8.8

ตารางที่ 4-1 ผลการทดลองจากการใช้ 1 คุณลักษณะ

จากผลการทดลอง จะเห็นได้ว่า จากการทดลองด้วยวิธีใดวิธีหนึ่งเพียงอย่างเดียว ได้แก่HOG, LDP, LBP, LHSV ผลการทดลองที่ได้ความแม่นยำสูงสุด คือวิธี LDP ซึ่งมีความแม่นยำอยู่ที่ 59.7 % และจะเห็นว่าวิธี LHSV ซึ่งใช้คุณสมบัติของสีเพียงอย่างเดียวมีความแม่นยำอยู่ที่ 43.8 %เท่านั้น

4.1.2 การแยกแยะด้วย 2 คุณลักษณะ

จากวิธีที่ใช้ในการทดสอบของการแยกแยะด้วย1 คุณลักษณะ จะสังเกตเห็นได้ว่า คุณลักษณะที่ใช้ ของวิธี HOG, LDP, LBP จะเป็นการใช้คุณลักษณะของเส้น หรือ ภาพ ที่เป็น ภาพสีเทา (gray scale) ส่วนวิธี LHSV จะใช้คุณลักษณะของสีเป็นหลัก ดังนั้นจึงคิดว่า หากนำคุณสมบัติของสี มาใช้ร่วม กับคุณลักษณะของเส้น ด้วยน่าจะทำให้ได้ผลการทดลองที่ดีขึ้นกว่าเดิม จึงทำการทดลองเพิ่มอีก 3 วิธี ได้แก่

- HOG + LHSV
- LDP + LHSV
- LBP + LHSV

สำหรับขั้นตอนการทดลองจะทำเช่นเดียวกับการแยกแยะด้วย 1 คุณลักษณะ ซึ่งจะได้ผลการทดลองดังตารางที่ 4-2

วิธีที่ใช้	%เฉลี่ยความแม่นยำจากการทดลอง	ส่วนเบี่ยงแบนมาตรฐาน
HOG	59.6	3.3
LBP	56.2	7.4
LDP	59.7	4.1
LHSV	43.8	8.8
HOG + LHSV	62.8	6.1
LDP + LHSV	59.3	5.1
LBP + LHSV	66.8	5.0

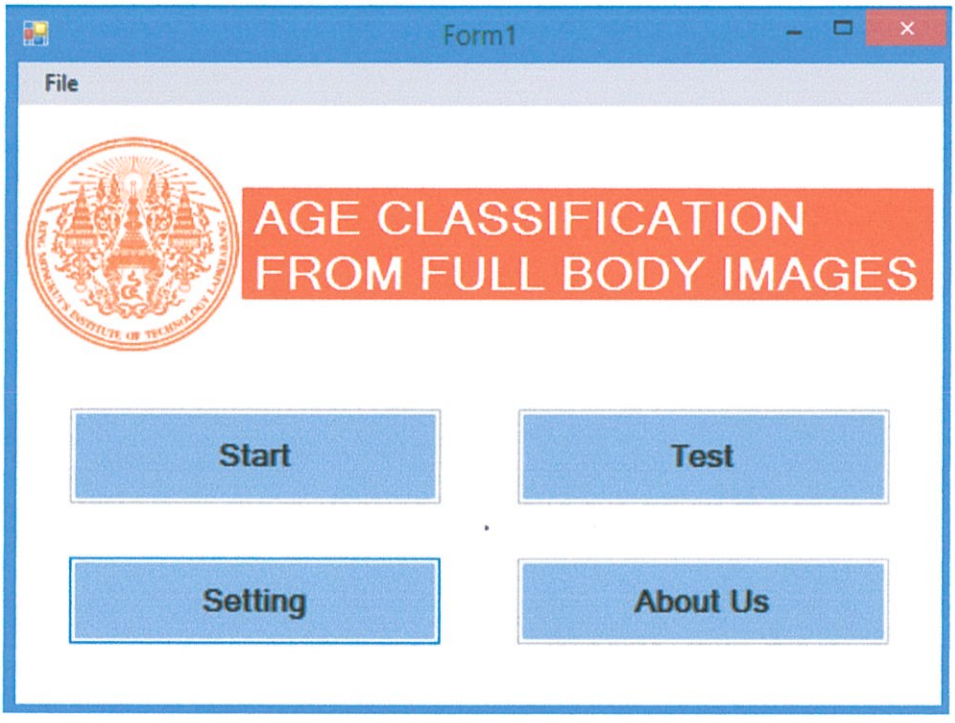
4-2 ผลการทดลองจากการใช้ 2 คุณลักษณะร่วมกัน

เมื่อทำการทดลองโดยใช้วิธี LHSV ร่วมกับวิธีอื่น จะเห็นว่าผลการทดลองมีความถูกต้องสูงขึ้น โดยวิธีที่มีความแม่นยำสูงสุดได้แก่วิธี LBP + LHSV ซึ่งได้ผลความถูกต้องอยู่ที่ 66.8 % ทั้งนี้ การที่วิธี LBP + LHSV ได้ความถูกต้องสูงขึ้นกว่าการใช้ LBP เพียงอย่างเดียว มาก เนื่องมาจากการใช้คุณลักษณะที่ต่างกันของวิธี LBP และ LHSV ทำให้ได้ vector ที่ต่างกัน ซึ่งทำให้ทั้ง 2 วิธีช่วยส่งเสริมกันทำให้สามารถสังเกตลักษณะต่างๆ ได้ดีขึ้นจึงทำให้ได้ความถูกต้องที่สูงขึ้น ส่วนวิธีอื่นที่ได้ความถูกต้อง ไม่สูงขึ้นมากนัก อาจเกิดจาก คุณลักษณะทีใช้ในการคิดคำนวณมีบางส่วนที่คล้ายกัน จึงได้ vector ที่ข้อมูลมีส่วนคล้ายกัน จึงไม่ช่วยให้ความถูกต้องสูงขึ้นมากนัก

4.2 ผลการพัฒนาโปรแกรม

ขั้นตอนการใช้งาน

<u>หน้าหลักของโปรแกรม</u>



ภาพที่ 4-1 หน้าหลักของโปรแกรม

หน้าแรกจะประกอบไปด้วย ปุ่มทั้งหมด 4 ปุ่ม ได้แก่ ปุ่ม Start, ปุ่ม Test, ปุ่ม Setting และ ปุ่ม About Us ดังภาพที่ 4-1 ซึ่ง

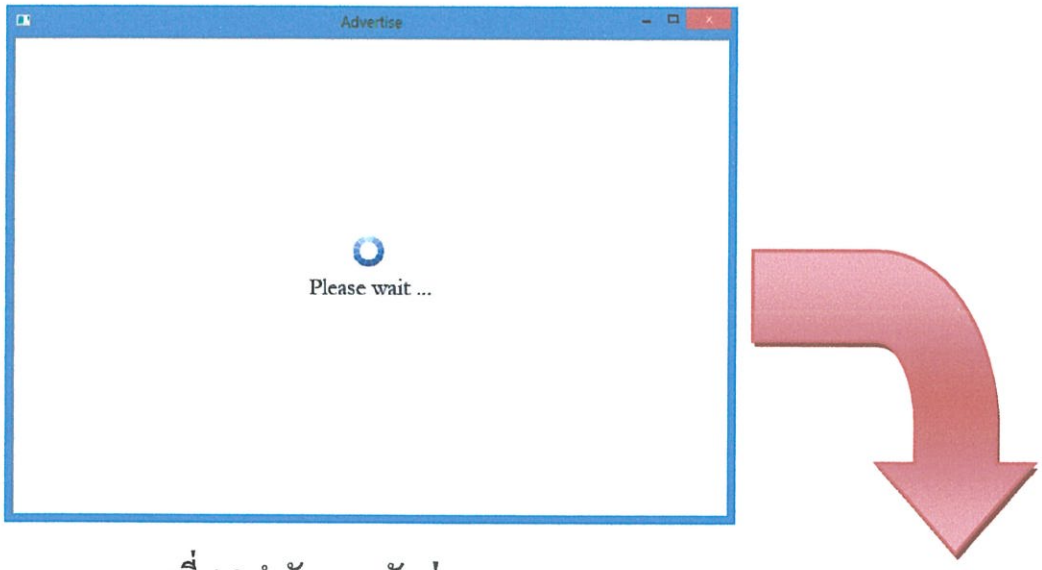
-ปุ่ม Start จะเป็นปุ่มที่ใช้เริ่มการทำงานของโปรแกรมในแบบ real time

-ปุ่ม Test จะเป็นปุ่มที่ใช้ทดสอบการทำงานของโปรแกรม

- ปุ่ม Setting จะเป็นปุ่มที่ใช้ในการตั้งค่าหรือเปลี่ยนโฆษณาสำหรับแต่ละช่วงอายุ

-ปุ่ม About Us จะเป็นปุ่มที่แสดงข้อมูลของผู้พัฒนาและเทคนิคที่ใช้ในการทำ

<u>หน้าการทำงานของโปรแกรมในแบบ real time</u>



ภาพที่ 4-2 กำลังตรวจจับช่วงอายุบุคล

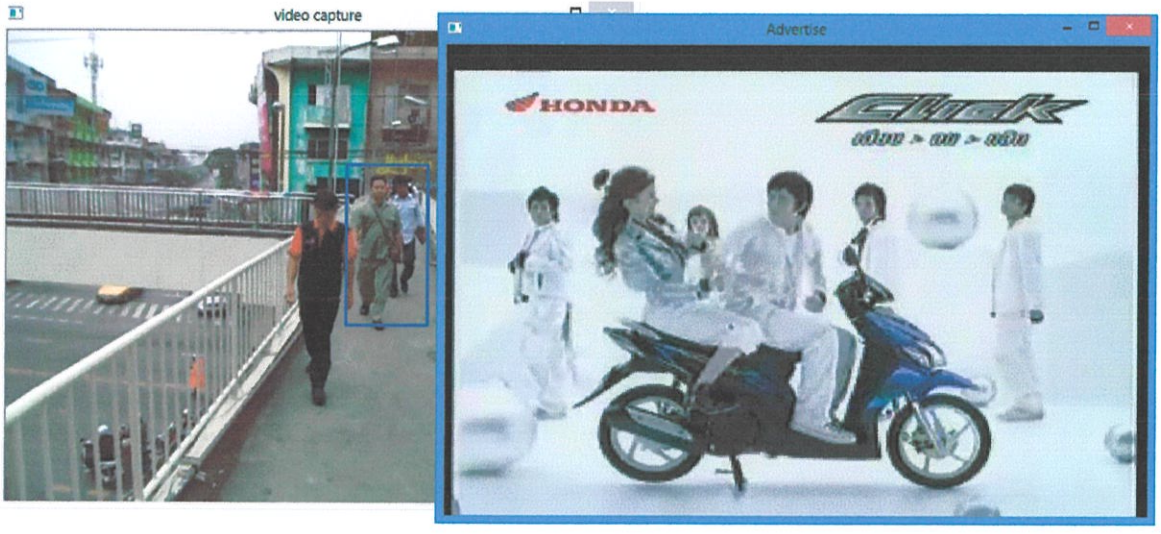
ภาพที่ 4-3 การแสดงของโฆษณา

เมื่อกดปุ่ม Start จะเข้าสู่การทำงานของโปรแกรม ซึ่งจะมีหน้าจอใหม่ขึ้นมาเป็นหน้าจอของการแสดงโฆษณา ดังภาพที่ 4-2 ซึ่งจะแสดงว่าโปรแกรมกำลังตรวจจับช่วงอายุของบุคล และเมื่อตรวจจับได้แล้ว ก็จะแสดงโฆษณาตามช่วงอายุของบุคคลตามที่ตรวจจับได้ ดังภาพที่ 4-3

<u>หน้าทดสอบการทำงานของโปรแกรม</u>

ภาพที่ 4-4 กำลังตรวจจับช่วงอายุบุคล

ภาพที่ 4-5 ตรวจจับช่วงอายุบุคลสำเร็จ



ภาพที่ 4-6 การแสดงของโฆษณาของหน้าทดสอบการทำงานของโปรแกรม

เมื่อกดปุ่ม Test จะเข้าสู่การทดสอบการทำงานของโปรแกรม ซึ่งจะมีหน้าจอใหม่ขึ้นมาเป็นหน้าจอของการแสดงโฆษณา ดังภาพที่ 4-4 ซึ่งกรอบสีขาวจะแสดงว่าโปรแกรมกำลังตรวจจับช่วงอายุของบุคล ซึ่งกรอบแต่ละสีจะระบุแต่ละช่วงอายุดังนี้

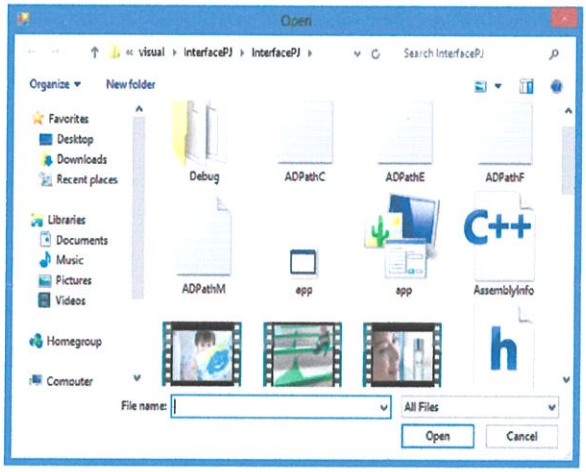
- สีขาว = กำลังตรวจจับช่วงอายุบุคคล
- สีเหลือง = วัยเด็ก
- สีน้ำเงิน = วัยรุ่น-ผู้ใหญ่ (ผู้ชาย)
- สีชมพู = วัยรุ่น-ผู้ใหญ่ (ผู้หญิง)

- สีเขียว = วัยสูงอายุ

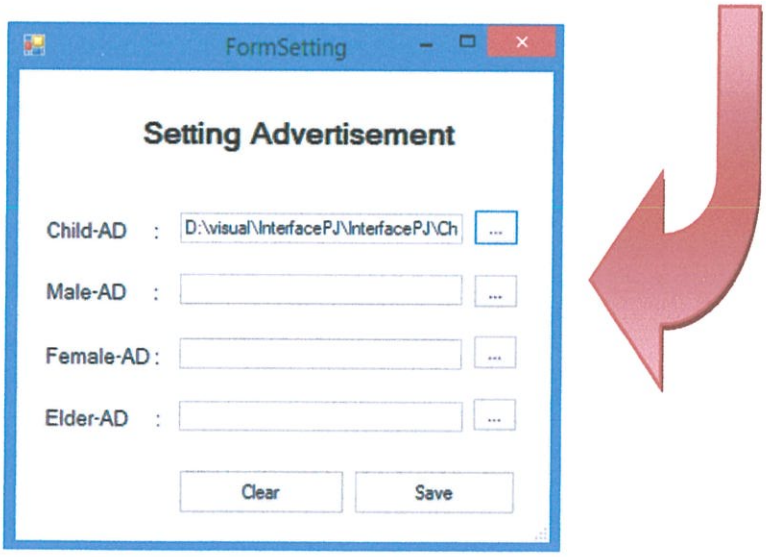
และเมื่อตรวจจับช่วงอายุบุคลได้แล้วดังภาพที่ 4-5 ก็จะแสดงโฆษณาตามช่วงอายุของบุคคลตามที่ตรวจจับได้ในหน้าจอใหม่ ดังภาพที่ 4-6

<u>หน้าการตั้งค่าหรือเปลี่ยนโฆษณาสำหรับแต่ละช่วงอายุ</u>

ภาพที่ 4-7 หน้าการตั้งค่าหรือเปลี่ยนโฆษณา



ภาพที่ 4-8 หน้าการค้นหาและเลือกโฆษณา

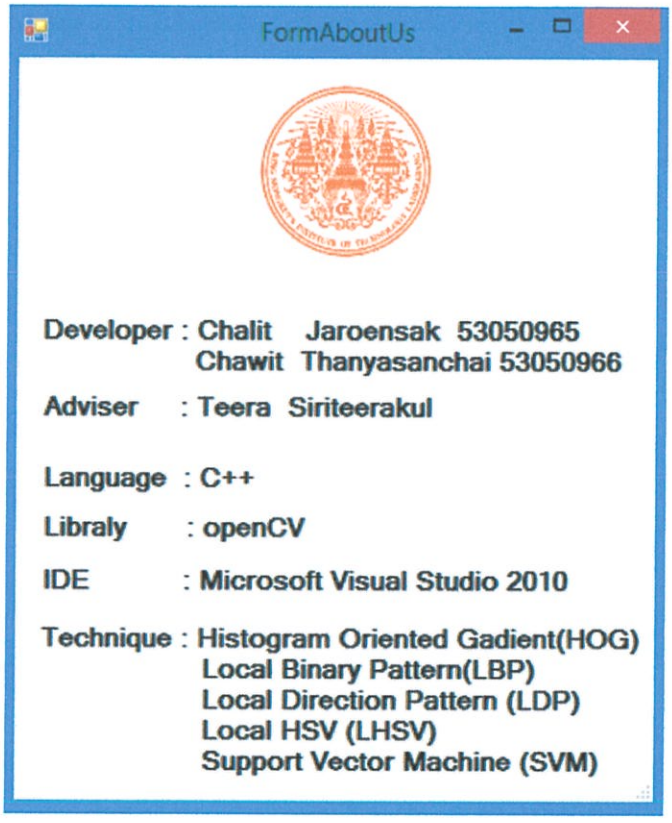


ภาพที่ 4-9 หน้าแสดงเมื่อตั้งค่าหรือเปลี่ยนโฆษณา

เมื่อกดปุ่ม Setting ก็จะเข้าสู่หน้าสำหรับการตั้งค่าหรือเปลี่ยนโฆษณาสำหรับแต่ละช่วงอายุของบุคคลดังภาพที่ 4-7 โดยจะกดปุ่ม ... เพื่อทำการค้นหาและเลือกโฆษณาที่ต้องการ ซึ่งเมื่อกดปุ่ม ... ก็จะปรากฏหน้าจอดังภาพที่ 4- 8 และเมื่อเลือกโฆษณาเสร็จแล้วก็จะได้ดังภาพที่ 4- 9

เมื่อเลือกโฆษณาเสร็จแล้วให้กีที่ปุ่ม Save เพื่อบันทึก หรือ ถ้าต้องการลบโฆษณาที่เลือกมาทั้งหมด แล้วต้องการเลือกโฆษณาใหม่ให้กดที่ปุ่ม Clear

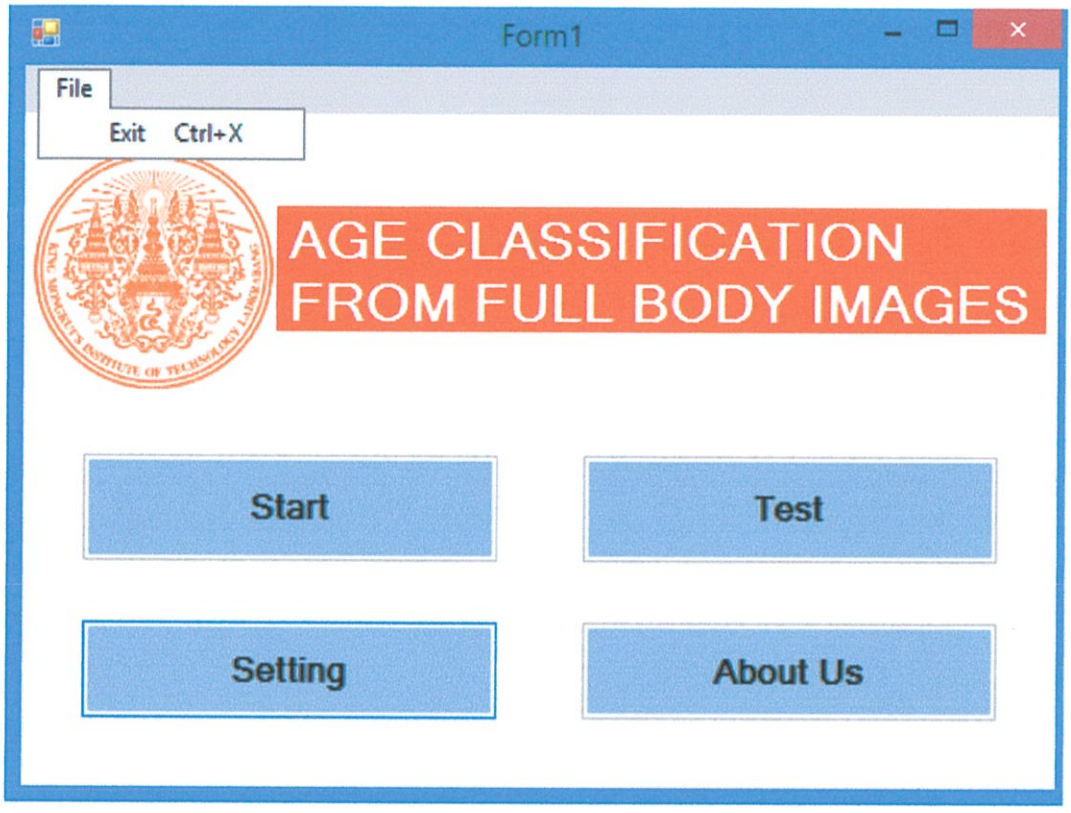
<u>หน้าแสดงข้อมูลของผู้พัฒนาและเทคนิคที่ใช้ในการทำ</u>



ภาพที่ 4-10 หน้าแสดงข้อมูลของผู้พัฒนาและเทคนิคที่ใช้ในการทำ

เมื่อกดปุ่ม About Us ก็จะเข้าสู่หน้าสำหรับแสดงข้อมูลของผู้พัฒนาและเทคนิคที่ใช้ในการทำดังภาพที่ 4-10

การออกจากโปรแกรม



ภาพที่ 4-11 การออกจากโปรแกรม

การออกจากโปรแกรมสามารถกดที่ปุ่ม ตามปกติได้เลย หรือจะเลือกที่ File->Exit หรือจะกดปุ่ม Ctrl+X ก็ได้

บทที่ 5

สรุปผล และข้อเสนอแนะ

5.1 สรุปผลการดำเนินงาน

จากการทำการทดลองเพื่อหาความถูกต้องแม่นยำในการแยกแยะ เพศ และ ช่วงอายุของบุคคล จะเห็นได้ว่าการทดลองด้วยวิธีต่างๆ ได้แก่ HOG, LBP, LDP, LHSV, HOG+LHSV, LDP+LHSV, LBP+LHSV ซึ่งแต่ละวิธีเหล่านี้ใช้คุณสมบัติที่ใช้ในการคิดคำนวณที่แตกต่างกัน ซึ่งทำให้ความถูกต้องแม่นยำที่ได้แตกต่างกันด้วย จากผลการทดลองวิธีที่ได้ความถูกต้องแม่นยำสูงสุดได้แก่ วิธี LBP + LHSV ซึ่งได้ความแม่นยำ 66.8 % ซึ่งไม่สูงมากนัก แต่เนื่องจาก วิธีการตรวจจับภาพแบบเต็มตัว ในปัจจุบัน ยังมีความแม่นยำไม่สูงมากนัก รวมทั้ง กลุ่มข้อมูลที่ใช้ในการเรียนรู้ถูกแบ่ง ออกเป็น 4 กลุ่มด้วยกัน ดังนั้น ความแม่นยำที่ 66.8 % จึงถือได้ว่า เป็นที่น่าพึงพอใจในระดับหนึ่ง

จากการพัฒนาโปรแกรมในการตรวจจับภาพคนแบบเต็มตัวเพื่อแบ่งแยกเพศและช่วงอายุ เพื่อใช้ในการส่งเสริมการโฆษณา ผลที่ได้ คือ

- โปรแกรมสามารถ ตรวจจับภาพคนและ แยกแยะ เพศ และช่วงอายุ ได้ในระดับหนึ่ง
- สามารถตรวจจับภาพได้แบบ realtime
- สามรถแสดง โฆษณาที่ได้ตั้งไว้ ตามช่วงอายุ และ เพศ ของบุคคลที่ตรวจจับได้
- สามารถปรับเปลี่ยน โฆษณา ที่จะแสดงสำหรับแต่ละเพศ และ ช่วงอายุได้

5.2 ปัญหาในการดำเนินการ

- การใช้วิธี HOG ในการตรวจจับภาพบุคคล สามารถตรวจจับได้ทั้งภาพคนหันหน้า หันหลัง หรือด้านข้าง แต่ชุดข้อมูล Data set ที่ใช้มีเพียงภาพด้านหน้าเพียงอย่างเดียว ทำให้ ถ้าภาพที่ได้จากการตรวจจับเป็นภาพด้านข้างหรือด้านหลัง จะได้ผลที่ผิดเพี้ยนไป
- ในการตรวจจับ ถ้าบริเวณนั้นมีแสงสว่างไม่เพียงพอจะทำให้ ตรวจจับได้ยาก หรือผลที่ได้คลาดเคลื่อนได้ง่าย
- เนื่องจากชุดข้อมูลเป็นชุดข้อมูลที่เกิดจากการเก็บข้อมูลเอง จึงเป็นการยากที่จะทำการเก็บข้อมูลของคนให้ได้ในปริมาณที่สูงเพียงพอในทุกช่วงอายุ
- สำหรับช่วงอายุ ของผู้ใหญ่และ ผู้สูงอายุ แล้ว มีลักษณะที่คล้ายกัน ทำให้แยกได้ยาก

5.3 ข้อเสนอแนะ

ในการตรวจจับภาพบุคคลและแยกแยะเพศและช่วงอายุ รายละเอียดต่าง ๆ เป็นสิ่งที่มองข้ามไม่ได้แม้จะเป็นรายละเอียดเล็กน้อยก็ตามก็อาจทำให้ผลที่ได้คลาดเคลื่อนไปจากที่ต้องการได้ โดยส่วนที่ควรคำนึงถึงเป็นสำคัญได้แก่

- ปริมาณของชุดข้อมูลควรมีมากเพียงพอ ในแต่ละกลุ่มข้อมูล
- การคัดแยกกลุ่มข้อมูลแต่ละกลุ่มควรแยกออกจากกันอย่างชัดเจน
- สภาพแวดล้อมมีส่วนสำคัญมาก ในการตรวจจับและแสดงผล เช่น แสงสว่าง และสถานที่ ที่ต่างกันระหว่างตรวจจับ กับ ภาพชุดข้อมูลที่ใช้ในการเรียนรู้ ก็มีส่วนทำให้ผลที่ได้ผิดเพี้ยน
- ภาพที่ใช้ในการเรียนรู้ของเครื่อง และ ภาพที่ได้จากกล้องที่ใช้ในการตรวจจับ หากมีความละเอียดต่างกันมาก ก็สามารถส่งผลให้ผลที่ได้ผิดพลาดได้ เช่นกัน

เอกสารอ้างอิง

[1] Matthew Collins, Jianguo Zhang, Paul Miller,Hongbin Wang ,"Full body image feature representations for gender profiling" , The Institute of Electronics, Communications and Information Technology (ECIT) Queens University Belfast

[2] K. van de Sande, T. Gevers, and C. Snoek.Evaluation of color descriptors for object and scene recognition. CVPR 2008. IEEE Conference on, pages 1–8, 2008.

[3] Yasushi Makihara, Hidetoshi Mannami, and Yasushi Yagi,"Gait Analysis of Gender and Age Using a Large-Scale Multi-View Gait Database", Osaka University, 8-1 Mihogaoka, Ibaraki, Osaka, 567-0047, Japa

[4] Navneet Dalal and Bill Triggs,"Histograms of Oriented Gradients for Human Detection" ,INRIA Rhone-Alps,655 avenue de l'Europe,Montbonnot 38334,France

[5] Porikli, F. and Yilmaz, A. "Object Detection & Tracking" , MITSUBISHI ELECTRIC RESEARCH LABORATORIES

[6] Hasanul Kabir, Taskeed Jabid, and Oksam Chae. "Local Directional Pattern Variance (LDPv): A Robust Feature Descriptor for Facial Expression Recognition" , Department of Computer Engineering, Kyung Hee University, South Korea.