

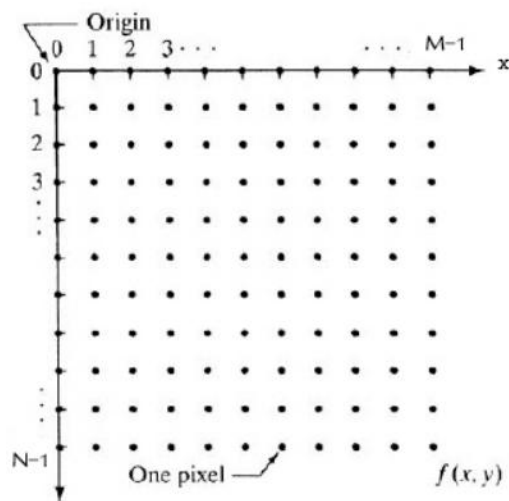
บทที่ 2

ทฤษฎีที่เกี่ยวข้องกับการดำเนินงาน

โครงการนี้จะเน้นเรื่องการประมวลผลภาพดิจิทัลสำหรับใช้ในการตรวจจับโมเดล ส่วนท้ายรถยนต์ที่นิยมใช้ในประเทศไทยโดยการนำอัลกอริทึม Scale Invariant Feature Transform (SIFT) ทฤษฎีและหลักการต่าง ๆ ที่จะกล่าวต่อไปนี้เป็นประยุกต์ใช้ และทฤษฎีของการประมวลผลภาพดิจิทัลที่เกี่ยวข้องรวมถึงทฤษฎีที่นำมาใช้กับระบบตรวจจับโมเดลส่วนท้ายรถยนต์แบบอัตโนมัติที่จะพัฒนาขึ้นในโครงการนี้และสามารถนำไปพัฒนาต่อในอนาคต

2.1 ภาพและความหมายของพิกเซล

2.1.1 พิกเซล (Pixel) คือ ความเข้มแสงที่รวมกันทำให้เกิดเป็นภาพ ภาพหนึ่ง ๆ จะประกอบด้วยพิกเซลมากมาย ซึ่งภาพแต่ละภาพที่สร้างขึ้นจะมีความหนาแน่นของพิกเซลเหล่านี้แตกต่างกันไป ความหนาแน่นนี้เป็นตัวบอกถึงความละเอียด (Resolution) ของภาพซึ่งมีหน่วยเป็น ppi (Pixel Per Inch) คือจำนวนพิกเซลต่อนิ้ว ซึ่งโดยทั่วไปถือว่าภาพที่มีความละเอียดสูงหรือคุณภาพดีจะมีความละเอียด 300 x 300 ppi ขึ้นไป ค่า ppi ยิ่งสูงขึ้น ภาพก็จะมี ความละเอียดและคมชัดมากขึ้น



รูปที่ 2.1 ตำแหน่งของพิกเซล

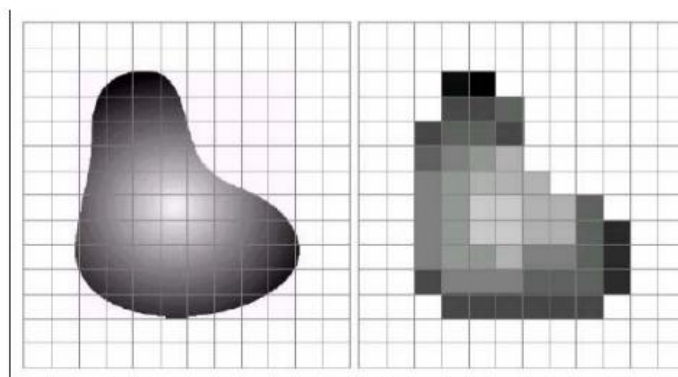
$N =$ จำนวนพิกเซลที่มากที่สุดในแกน Y

$M =$ จำนวนพิกเซลที่มากที่สุดในแกน X

2.2 การประมาณผลภาพเชิงตัวเลข (Digital Image Processing)

2.2.1 การแทนภาพด้วยภาพแบบดิจิทัล ภาพแบบดิจิทัล (Digital Image) เป็นภาพที่ถูกแปลงมาจากอนาลอก ให้อยู่ในรูปของตัวเลขโดยภาพอนาลอกถูกแบ่งเป็นพื้นที่สี่เหลี่ยมเล็ก ๆ ที่เรียกว่าพิกเซล ในแต่ละพิกเซลจะถูกระบุตำแหน่งด้วยคู่ออติเนต x,y และค่าระดับความเข้มของแสงของพิกเซลนั้น ๆ โดยเราสามารถแปลงภาพเป็นแบบดิจิทัลโดยมีขั้นตอนและวิธีการดังนี้

เมื่อเรานำสัญญาณอนาลอกที่ต้องการประมวลผลผ่านส่วนที่เรียกว่าดิจิทัลไรเซอร์ (Digitizer) ซึ่งจะทำหน้าที่ในการแปลงสัญญาณอนาลอกให้เป็นสัญญาณดิจิทัล จากนั้นทำการควอนไทซ์ (Quantizing) เพื่อแปลงค่าความเข้มของแสงให้เป็นตัวเลข พังก์ชันของภาพ $f(x,y)$ จะถูกทำให้เป็นสัญญาณไม่ต่อเนื่องทั้งระนาบของภาพซึ่งเรียกว่า การสุ่มภาพ (Image Sampling) ของฟังก์ชันที่ได้เรียกว่า การควอนไทซ์ ระดับความเข้มของแสง (Greasy Level Quantization) ก็จะได้ข้อมูลที่เป็นดิจิทัลดังภาพที่ 2.2



รูปที่ 2.2 การทำ Sampling และ Quantization

2.2.2 ลักษณะการจัดเก็บรูปภาพแบบดิจิทัล

โดยทั่วไปแล้วภาพจะมีความเข้มตั้งแต่ 2 ระดับขึ้นไป แต่ที่นิยมใช้กันมาก คือ ค่าระดับความเข้มของพิกเซลที่เท่ากับ 256 ระดับ ซึ่งจะทำให้ค่าของพิกเซลอยู่ในช่วง (0-255) โดยใช้เนื้อที่ในการเก็บข้อมูลขนาด 1 ไบต์หรือ 8 บิต สำหรับข้อมูล 1 พิกเซล (256) ในกรณีที่ต้องการภาพที่มีความเข้มสูงอาจจะต้องการจำนวนบิตสำหรับการเก็บข้อมูลมากกว่า 8 บิต คืออาจจะเป็น 16 หรือ 24 บิตโดยจะแยกความแตกต่างของภาพแต่ละประเภทให้เห็นอย่างชัดเจนได้ดังนี้

2.2.2.1 ภาพ 2 ระดับ คือ มีพิกเซลสีขาวกับสีดำเท่านั้น โดยแต่ละพิกเซลจะมีขนาดของข้อมูลเท่ากับ 1 บิต

2.2.2.2 ภาพ 16 ระดับ คือ ในแต่ละพิกเซลจะมีขนาดของข้อมูล 8 บิต ซึ่งสามารถแสดงภาพได้ความเข้มถึง 256 ระดับ

2.2.2.3 ภาพ 256 ระดับ คือ ในแต่ละพิกเซลจะมีขนาดของข้อมูล 8 บิต ซึ่งทำให้สามารถแสดงภาพได้ความเข้มถึง 256 ระดับ

2.2.2.4 ภาพทิวทัศน์ (True Color) คือในแต่ละพิกเซลจะมีขนาดของข้อมูล 24 บิตทำให้สามารถแสดงภาพออกมาได้เฉพาะภาพสีเท่านั้น ไม่สามารถแสดงเป็นภาพขาวดำได้

โดยทั่วไปวิธีการประมวลผลภาพเชิงตัวเลขที่ทำให้คอมพิวเตอร์สามารถรู้วัตถุในภาพได้นั้นแบ่งออกได้เป็นสองระดับด้วยกันคือ การประมวลผลภาพในระดับต่ำ (Low-Level Image) และการประมวลผลภาพในระดับสูง (High-Level Image Processing) การประมวลผลภาพในระดับต่ำจะเป็นการประมวลผลเชิงตัวเลขเกือบทั้งหมด เพื่อหาตัวแปรต่าง ๆ มาอธิบายข้อมูลภาพ โดยมีจุดประสงค์ที่จะนำตัวแปรเหล่านั้นไปใช้ในการประมวลผลภาพระดับสูงต่อไปโดยทั่วไปแล้วการประมวลผลภาพระดับต่ำจะประกอบด้วย การประมวลผลภาพก่อน (Preprocessing) เช่น การกำจัดสัญญาณรบกวน หรือการทำให้ภาพคมชัดการหาของภาพ เป็นต้น

การประมวลผลระดับสูงเป็นการนำผลลัพธ์ หรือสัญลักษณ์ที่ได้จากการประมวลผลภาพระดับต่ำมาตีความหรือประมวลผลเพื่อให้คอมพิวเตอร์สามารถรู้จักและเข้าใจภาพได้ สำหรับความแตกต่างของการประมวลผลภาพ ทั้ง 2 ประเภทนั้นคือ การประมวลผลภาพระดับต่ำจะใช้ค่าความสว่างของจุดภาพ (พิกเซล) ส่วนการประมวลผลภาพระดับสูงนั้นข้อมูลภาพที่นำมาประมวลผลจะถูกแสดงในรูปสัญลักษณ์ ซึ่งสัญลักษณ์เหล่านี้จะแสดงถึงสิ่งต่าง ๆ ที่อยู่ในภาพเช่น ขนาดของวัตถุรูปร่าง และความสัมพันธ์กันระหว่างวัตถุภาพ

2.3 การสร้างภาพไบนารี

การสร้างภาพไบนารีสามารถทำได้โดยใช้เทคนิคการทำเทรชโฮล (Thresholding Technique) โดยพิจารณาว่าพิกเซลสีขาวหรือสีดำ จะกระทำโดยการเปรียบเทียบระหว่างพิกเซลของภาพเริ่มต้นกับค่าคงที่หนึ่งเรียกว่า ค่าเทรชโฮล (Threshold Value) เทคนิคนี้ใช้กันมากในกรณีที่ข้อมูลภาพมีลักษณะที่ต่างกันระหว่างวัตถุ (Object) และพื้นหลัง (Background) โดยค่าของพิกเซลของภาพใด ๆ ที่มีค่ามากกว่าหรือเท่ากับค่าเทรชโฮลจะถูกเปลี่ยนเป็น 0 (จุดดำ) และพิกเซลของภาพที่มีน้อยกว่าหรือเท่ากับค่าเทรชโฮลจะถูกเปลี่ยนเป็น 1 (จุดขาว) เป็นต้น

ในการสร้างภาพไบนารีโดยใช้เทคนิคเทรชโฮลเพื่อให้ได้ผลลัพธ์ที่เหมาะสมและคมชัด สิ่งที่สำคัญที่สุดคือ ค่าเทรชโฮล เนื่องจากถ้าเลือกค่าเทรชโฮลที่ไม่เหมาะสม (ค่าเทรชโฮลที่มีค่าน้อยเกินไปหรือมากเกินไป) ภาพที่ได้อาจจะสว่างน้อยเกินไป หรือสว่างมากเกินไป หรือภาพที่ได้มีสิ่งรบกวน(Noise) เกิดขึ้น อันเป็นผลทำให้ภาพที่ได้ไม่สวยงามเท่าที่ควร ดังนั้นปัญหาของการสร้างภาพไบนารี ซึ่งมีวิธีคำนวณหาค่าเทรชโฮลโดยกำหนดค่าล่วงหน้า (Pre-assigned Threshold Value) การหาค่าเทรชโฮลจากค่ากลาง (Mid-Range Threshold Value) ซึ่งแต่ละวิธีอธิบายได้ดังนี้

2.3.1 การหาค่าเทรชโฮลโดยกำหนดค่าล่วงหน้าเป็นการกำหนดค่าเทรชโฮลโดยการกำหนดเองจากผู้ใช้งาน ซึ่งการกำหนดนี้จะขึ้นอยู่กับประสบการณ์ของผู้ใช้คนนั้น ๆ โดยการเลือกค่าคงที่ค่าหนึ่ง ซึ่งเรียกว่า ค่าเทรชโฮล โดยค่าที่เลือกมานี้จะเป็นค่าที่อยู่ระหว่างค่าต่ำและค่าสูงสุดของระดับความเข้มแสงของภาพ เช่นภาพอินพุทมีระดับความเข้มแสง 256 ระดับ จะมีค่าได้ตั้งแต่ 0-255 เมื่อเลือกค่าเทรชโฮลได้แล้วสามารถสร้างภาพไบนารีได้

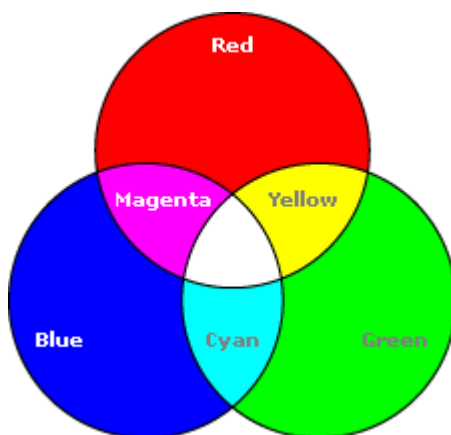
2.3.2 การหาค่าเทรชโฮลจากค่ากลาง เป็นการหาเทรชโฮลที่แตกต่างจากการหาค่าเทรชโฮลวิธีแรก สำหรับวิธีนี้จะเป็นการคำนวณหาค่าเทรชโฮลโดยอัตโนมัติโดยไม่ต้องให้ผู้ใช้เป็นผู้กำหนดโดยการหาค่าเทรชโฮลนี้ใช้วิธีทางสถิติในเรื่องการหาค่ากลางหรือค่าเฉลี่ย (Mean) ค่าเทรชโฮลที่คำนวณได้จากค่ากึ่งกลางที่อยู่ระหว่างค่าระดับความเข้มสูงสุด (Maximum Level) และระดับความเข้มต่ำสุด (Minimum Level) ของภาพ เมื่อการคำนวณค่าเทรชโฮลได้แล้วก็สามารถสร้างภาพไบนารีได้โดยนำค่าเทรชโฮลที่ได้มาใช้

2.4 แบบจำลองสี (Color Model)

แบบจำลองสี (Color Model) เป็นสิ่งที่ใช้อ้างอิงถึงสีต่าง ๆ สำหรับคอมพิวเตอร์แล้วเราจะไม่ใช้แบบจำลองที่เป็น Analytical Model เหมือนกับที่ใช้ในทางวิทยาศาสตร์ ซึ่งใช้วิธีการวัดซึ่งอยู่ในรูปของพลังงานช่วงของสเปกตรัม (Spectrum) แต่จะเป็น Empirical Model ที่ได้รับสัมพันธ์ของค่าที่ใช้อ้างอิงกับสีใด ๆ จากการทดลองที่เป็นการศึกษาแบบ Psychophysical ที่มีการรับรู้ของมนุษย์เข้ามาเกี่ยวข้อง

แบบจำลองสีมีหลายแบบด้วยกัน เช่น แบบจำลองสี RGB แบบจำลองสี CMY แบบจำลองสี CMYK แบบจำลองสี HSV แบบจำลองสี HIS แบบจำลองสี HLS แบบจำลองสี YIQ และแบบจำลองสี YUV แบบจำลอง YcbCR เป็นต้น

2.4.1 แบบจำลองสี RGB (RGB Color Model) เป็นแบบจำลองที่เฉพาะเจาะจงกับจอภาพคอมพิวเตอร์ เนื่องจาก RGB Model ได้ทำการสร้างสีต่าง ๆ ขึ้นโดยการใช้แหล่งกำเนิดแสงจำนวนสามสี ได้แก่ สีแดง (Red) สีเขียว (Green) และสีน้ำเงิน (Blue) ที่เกิดจากการเรียงแสงที่มีคุณสมบัติที่แตกต่างกันตามลำดับ ซึ่งแสงทั้งสามสีจะไม่เท่ากันในแต่ละอุปกรณ์ นอกเสียจากว่ามีคุณสมบัติของสารเรืองแสงและการตั้งค่าจอภาพ และสภาพแวดล้อมที่จอภาพคอมพิวเตอร์เหมือนกันทุกประการ ซึ่งโดยปกติแล้วจะมีค่าที่แตกต่างกันออกไปตามภาพที่ 2.3



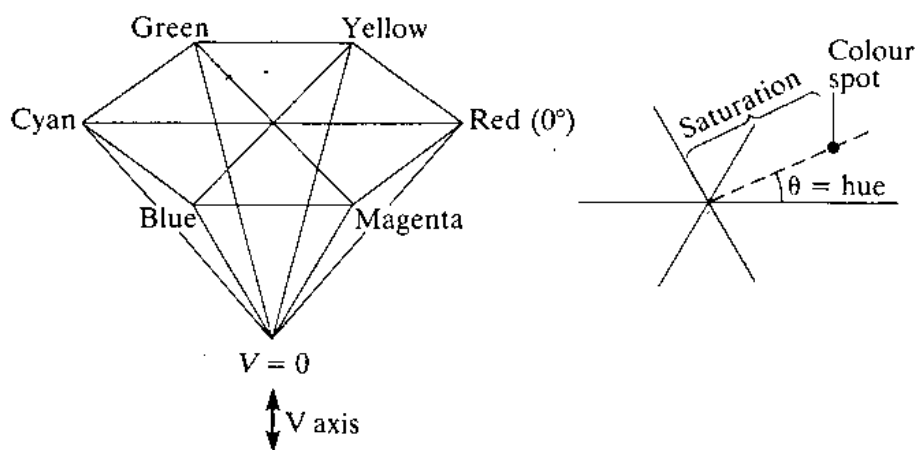
รูปที่ 2.3 แบบจำลองสี RGB

จำลองสี RGB ประกอบด้วยข้อมูลจำนวนสามส่วนคือ ค่า Intensity ของสีทั้งสามซึ่งได้แก่สีแดง สีเขียว และสีน้ำเงิน

2.4.2 แบบจำลองสี HSV (HSV Color Model) เป็นแบบจำลองสีที่ถูกสร้างขึ้นเพื่อเป็นทางเลือกโดยแบบจำลองสี HSV แสดงในภาพที่ 2.4 จะให้ความหมายที่ดีกว่าเมื่อกล่าวถึง สีต่างๆ ในเชิงศิลปะ เช่น เมื่อพูดถึงสีเหลืองในทางศิลปะจะมีความแตกต่างกัน เมื่อพิจารณาสีเหลืองอ่อน สีเหลืองแก่ หรือสีน้ำตาลว่ามีความแตกต่างกันอย่างไร จะพบว่าทุกสี คือสีเหลืองนั่นเอง ที่มีระดับความเข้มหรือมีความอิ่มตัวที่แตกต่างกัน ดังนั้นสีในแบบจำลองสี HSV จึงให้ความรู้สึกที่เข้าใจได้มากกว่าสำหรับมนุษย์ซึ่งจำลองสี HSV ประกอบด้วยสามส่วนคือ

H หมายถึง Hue หรือสีที่มีค่าที่แตกต่างออกไปตามความถี่ของแสง เช่น แดง เหลือง เขียว น้ำเงิน หรือ ม่วง เป็นต้น

S หมายถึง Saturation หรือความอิ่มตัวของ Hue นั้นๆ เช่น สีแดง และสีชมพูคือ สีแดง เพียงแต่สีชมพูมีความอิ่มตัวน้อยกว่าภาพที่ 2.5 แบบจำลองสี HSV หมายถึง Value หรือค่าความสว่างของสี โดยที่ค่า Value ต่ำสุดหมายถึง สีดำ ไม่ว่าจะ Hue หรือ Saturation เท่าใด และค่า Value สูงสุดหมายถึง สีขาว ซึ่งเป็นสีที่สว่างที่สุดของ Hue และ Saturation นั้นๆ เช่น Hue ใด ๆ มีค่า Saturation มีค่าเท่ากับ 0 เมื่อ Value สูงที่สุดคือสีเหลืองและ Value ต่ำสุดคือสีดำ



รูปที่ 2.4 แบบจำลองสี HSV

2.5 การแยกลักษณะเฉพาะของภาพ (Image Feature Extraction)

การแยกลักษณะเฉพาะของภาพเป็นการแยกหรือสกัดเอาข้อมูลที่สำคัของภาพออกมา ซึ่งลักษณะเฉพาะของภาพเป็นคุณสมบัติที่สามารถหาได้โดยใช้ขั้นตอนวิธีการประมวลผลภาพ (Image Processing) โดยที่ลักษณะเฉพาะพื้นฐานของภาพประกอบด้วย 3 ส่วนคือสีรูปร่างและพื้นผิว

- สี (Color) เป็นลักษณะเฉพาะของภาพที่มีบทบาทสำคัญในระบบค้นคืนภาพเช่นฮิสโตแกรมสีซึ่งเป็นลักษณะเฉพาะของสีที่ถูกลำมาใช้บ่อยๆเนื่องจากสีเป็นสิ่งที่สามารถมองเห็นได้ง่ายและเป็นสิ่งแรกที่สามารถสังเกตเห็นได้จากการมองภาพนอกจากนี้สียังสามารถใช้ในการแยกแยะกลุ่มของภาพออกตามเนื้อหาได้เป็นอย่างดีเช่นสีฟ้าของน้ำทะเลสีแดงของดอกไม้สีเขียวของต้นไม้ เป็นต้น

- รูปร่าง (Shape) เป็นลักษณะเฉพาะของภาพที่ใช้อธิบายถึงรูปร่างและลักษณะรวมถึงขนาดของวัตถุภายในภาพซึ่งทำให้สามารถแยกวัตถุออกจากพื้นหลังหรือแยกแยะระหว่างวัตถุที่มีรูปร่างแตกต่างกันออกจากกันได้

- พื้นผิว (Texture) เป็นลักษณะเฉพาะที่ใช้อธิบายความหยาบความละเอียดหรือความซับซ้อนของวัตถุภายในภาพซึ่งแต่ละภาพอาจจะประกอบด้วยวัตถุที่มีลักษณะพื้นผิวที่แตกต่างกันออกไปการวิเคราะห์พื้นผิวจะช่วยให้สามารถแยกแยะความแตกต่างของวัตถุได้ดียิ่งขึ้นการค้นคืนภาพที่ใช้พื้นผิวเป็นลักษณะเฉพาะของภาพส่วนใหญ่จะถูกนำไปใช้ในการค้นหาภาพจากกลุ่มภาพพื้นผิวเช่นชุดภาพพื้นผิวของหินชุดภาพพื้นผิวของใบไม้ เป็นต้น

2.6 การสร้างดัชนีภาพ (Image Indexing)

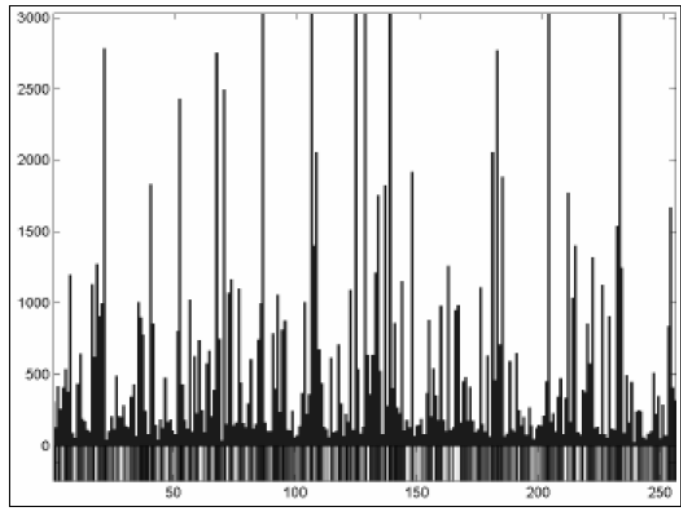
การสร้างดัชนีภาพมีด้วยกันหลากหลายวิธีซึ่งในที่นี้ขอยกตัวอย่างวิธีการนิยมและได้การยอมรับมา 4 วิธี คือ

2.6.1 ฮิสโตแกรมสี

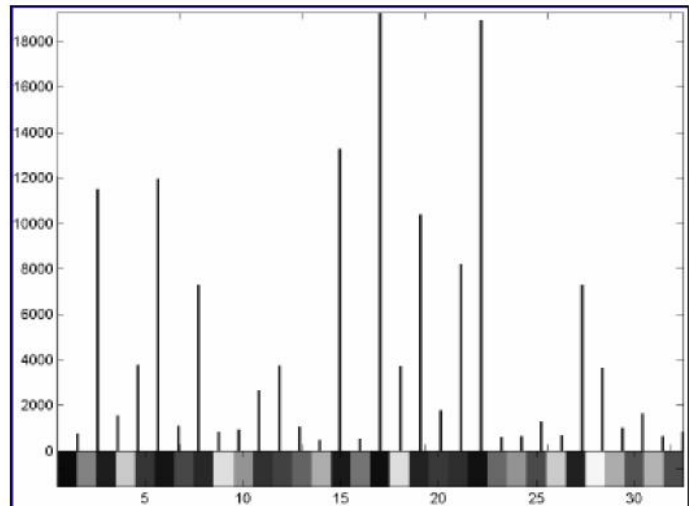
ฮิสโตแกรมสีเป็นกราฟแสดงความถี่หรือจำนวนจุดภาพที่ระดับค่าสีแต่ละค่าโดยแกนนอนแทนระดับค่าสีต่างๆ ส่วนแกนตั้งแทนจำนวนจุดภาพที่ระดับค่าสีนั้นๆเป็นลักษณะเฉพาะของ

ภาพในลักษณะของการวัดการแจกแจงของสีในภาพซึ่งสามารถพิจารณาได้จากภาพตัวอย่างที่ 2-6 และ 2-7

สำหรับภาพสี RGB ขนาด 24 บิตต่อจุดภาพแต่ละ Channel คือ แดง เขียวและน้ำเงิน นั้นจะมีค่าความสว่าง 8 บิตหรือ 256 ระดับโดยมีค่าตั้งแต่ 0-255 ดังนั้นจุดภาพแต่ละจุดจึงสามารถแสดงสีได้มากถึง 16.7 ล้านสี



รูปที่ 2.5 ตัวอย่างภาพและฮิสโตแกรมของภาพที่มี 256 กลุ่มสี



รูปที่ 2.6 ตัวอย่างภาพและฮิสโตแกรมของภาพที่มี 32 กลุ่มสี

ในการคำนวณค่าฮิสโตแกรมสีภาพแต่ละภาพจะถูกควอนไทซ์สีภายในภาพเพื่อลดมิติของเวกเตอร์และลดความซับซ้อนในการคำนวณลงโดยการแบ่งกลุ่มสีออกเป็น m ถังสี (Bins) ซึ่งส่วนใหญ่นิยมใช้ 32, 64 หรือ 256 ถังสีเนื่องจากการแยกแยะความแตกต่างของระดับค่าสีของสายตามนุษย์มีความละเอียดไม่มากนักกำหนดให้ภาพ I มีขนาด $n_1 \times n_2$ จุดภาพและ H_{c_i} แทนจำนวนจุดภาพที่มีสี c_i ของภาพ I ดังนั้นสามารถคำนวณฮิสโตแกรมสีได้ดังสมการที่ 2-1

$$h_{c_i}(I) = \frac{H_{c_i}}{n_1 \times n_2} \quad (2-1)$$

โดยที่ $h_{c_i}(I)$ คือฮิสโตแกรมของสี c_i ของภาพ I

M คือจำนวนสีภายในภาพหลังการควอนไทซ์สี

n_1 คือความกว้างของภาพ

n_2 คือความยาวของภาพ

การหารด้วย $n_1 \times n_2$ หรือจำนวนจุดภาพทั้งหมดภายในภาพมีวัตถุประสงค์เพื่อทำให้ค่าฮิสโตแกรมสีเป็นบรรทัดฐานและเพื่อให้สามารถนำฮิสโตแกรมสีของภาพซึ่งมีขนาดแตกต่างกันมาเปรียบเทียบกันได้

ดังนั้นจะได้เวกเตอร์แทนฮิสโตแกรมสีของภาพ I ดังสมการที่ 2-2

$$H(I) = (h_{c_1}, h_{c_2}, \dots, h_{c_m}) \quad (2-2)$$

ฮิสโตแกรมสีเป็นลักษณะเฉพาะทางสีของภาพที่นิยมนำมาใช้ในระบบการค้นคืนภาพส่วนใหญ่เนื่องจากสามารถคำนวณได้ง่ายและรวดเร็วแต่มีข้อเสียคือมีเฉพาะข้อมูลเกี่ยวกับการกระจายของสีภายในภาพเท่านั้นไม่มีข้อมูลเชิงตำแหน่ง (Spatial Information)

2.6.2 Color Coherence Vector (CCV)

การกำหนดการเชื่อมติดกันของสี (Color Coherence) ในลักษณะที่เป็นระดับซึ่งจุดภาพของสีนั้นๆ เป็นสมาชิกของบริเวณขนาดใหญ่ที่มีสีเหมือนกันเรียกบริเวณสำคัญนี้ว่าบริเวณที่เชื่อมติดกัน (Coherent Regions) มีความสำคัญในการแสดงลักษณะเฉพาะของภาพตัวอย่างเช่นภาพ 2 ภาพดังแสดงในรูปที่ 2.7 มีค่าฮิสโตแกรมของสีเหมือนกันทั้งคู่ ที่ภาพทั้งสองมีลักษณะที่

แตกต่างกันจะเห็นว่าสีสีแดงปรากฏอยู่ทั้ง 2 ภาพในปริมาณที่ใกล้เคียงกันภาพทางด้านซ้ายมีจุดภาพที่เป็นสีแดง (จากดอกไม้) กระจายอยู่อย่างกว้างๆ แต่ในขณะที่ภาพทางด้านขวามีจุดภาพที่เป็นสีแดง (จากเสื้อของนักกอล์ฟ) รวมกลุ่มอยู่ด้วยกันจากปัญหาดังกล่าว Pass และ Zabih จึงได้คิดค้นลักษณะเฉพาะของภาพรูปแบบใหม่ขึ้นมาคือ Color Coherence Vector หรือ CCV



รูปที่ 2.7 ตัวอย่างภาพที่มีฮิสโตแกรมเหมือนกัน

CCV คือฮิสโตแกรมสีที่มีการปรับปรุงโดยการเพิ่มเติมข้อมูลเชิงตำแหน่งเข้ามาหลักการของ CCV ก็คือการแยกกลุ่มจุดภาพออกเป็น 2 กลุ่มคือจุดภาพที่เชื่อมติดกัน (Coherent Pixel) และจุดภาพที่ไม่เชื่อมติดกัน (Incoherent Pixel) โดยจุดภาพใดๆ จะเป็นจุดภาพที่เชื่อมติดกันก็ต่อเมื่อองค์ประกอบที่เชื่อมติดกัน (Connected Component) ของจุดภาพนั้นมีค่ามากกว่าค่าคงที่ค่าหนึ่ง (π) ที่ได้กำหนดไว้ซึ่งในงานวิจัยของ Pass และ Zabih กำหนดให้มีค่าโดยประมาณเท่ากับ 1% ของจำนวนจุดภาพทั้งหมดภายในภาพส่วนจุดภาพที่เหลือก็จะเป็นจุดภาพที่ไม่เชื่อมติดกัน CCV เป็นการนำเสนอการแยกกลุ่มสำหรับแต่ละสีในภาพเพื่อกำจัดจุดภาพที่ไม่ได้เป็นองค์ประกอบสำคัญของภาพทิ้งไปโดยที่ CCV สามารถแยกจุดภาพที่เชื่อมติดกันออกจากจุดภาพที่ไม่เชื่อมติดกันและป้องกันการจับคู่ที่ผิดพลาดซึ่งสิ่งนี้เป็นลักษณะเด่นที่ฮิสโตแกรมสีไม่สามารถทำได้

CCV ของภาพใดๆคือฮิสโตแกรมสีของจุดภาพในกลุ่มจุดภาพที่เชื่อมติดกันของภาพนั้น กำหนดให้ $H_{ccv}(I)$ คือเวกเตอร์ขนาด m มิติแทน CCV ของภาพ I ซึ่งคำนวณได้จากสมการที่ 2-3

$$H_{ccv}(I) = \langle (\alpha_1, \beta_1), (\alpha_2, \beta_2), \dots, (\alpha_m, \beta_m) \rangle \quad (2-3)$$

ข้อดีของ CCV เมื่อเปรียบเทียบกับฮิสโตแกรมสีคือมีประสิทธิภาพในการค้นคืนภาพมากกว่าเนื่องจากมีการกำจัดจุดภาพที่ไม่ได้เป็นองค์ประกอบสำคัญของภาพทิ้งไปแต่จะใช้เวลาในการประมวลผลมากกว่าฮิสโตแกรมสี

2.7 การจับคู่ (Matching)

กระบวนการจับคู่ภาพที่เหมือนกัน (Matching) เป็นเหมือนการสื่อสารกันระหว่างภาพและขั้นตอนการพัฒนาการจับคู่รูปภาพนั้นขึ้นอยู่กับการตอบสนองหลังจากการทำ Localization ถ้าหากผู้ใช้แสดงให้เห็นว่าการทำ localization นั้นไม่ถูกต้อง การเรียนรู้จะเกิดขึ้นโดยการใช้คุณสมบัติหรือลักษณะเด่นของภาพทั้งหมดที่มีอยู่สำหรับทำ localization และกำหนดตำแหน่งโดยผู้ใช้ การเรียนรู้การจับคู่ภาพเป็นกระบวนการเรียนรู้ที่มีการสืบทอด คือ การสร้างคลังคำศัพท์หรือ dictionary และการรวบรวมข้อมูลสำหรับการจำแนกจะเพิ่มมากขึ้น คือ จำนวนของคำที่ได้จากการเรียนรู้จะมีเพิ่มมากขึ้นเพียงเวลาแค่ไม่นาน

การสร้าง Dictionary จะขึ้นอยู่กับการเพิ่มของ nearest neighbor classifier สำหรับคำหรือคุณสมบัติใหม่ที่ไม่อยู่ใน dictionary ถ้าถูกนำเข้ามาในระบบ แล้วมีการจำคู่ภาพและพบว่ามีส่วนที่คล้ายกับข้อมูลอื่นที่อยู่ใน dictionary ในขั้นตอนนี้หากระยะห่างระหว่างคำกับลักษณะเด่นของภาพอยู่ต่ำกว่า ค่าเกณฑ์ที่กำหนด จะทำให้มีการรู้จำคำนั้นเกิดขึ้น คือระบบสามารถระบุได้ว่าภาพที่มีการจับคู่เป็นภาพรอยขีดหรือรู้นั้น แต่ถ้าค่าระยะห่างสูงกว่าเกณฑ์ที่กำหนด (ภาพที่จับคู่กันมีส่วนที่เหมือนกันน้อยอาจไม่ใช่ภาพเดียวกัน) คำหรือชิ้นส่วนของภาพนั้น จะถูกเพิ่มเข้าไปเป็นคุณสมบัติและกำหนดเป็นค่าเริ่มต้นใหม่ของข้อมูลหรือ dictionary และนำไปใช้ในการ match ภาพในครั้งต่อไปที่มี ภาพแบบนี้เข้ามา

การหาจุดเด่นและการ Matching ภาพนั้นจะต้องใช้เวลานานขึ้นมาก และอาจมีประสิทธิภาพที่ไม่ดี ดังนั้นภาพที่จะนำมาหาจุดเด่นจึงต้องมีการ ปรับขนาดซึ่งอาจผ่านการ แยกแยะรูปร่างมาก่อนแล้ว จะถูกนำมาแปลงขนาดให้เท่ากับขนาดของ template ที่เก็บไว้ใน database หลังจากนั้นจึงทำหมุนภาพให้มี orientation ที่สอดคล้องกับภาพใน template โดยอ้างอิงจากรูปร่างของรถยนต์เป็นหลัก

2.8 SIFT (Scale Invariant Feature Transform)

Scale invariant feature transform หรือ SIFT เป็นอัลกอริทึมหนึ่งในคอมพิวเตอร์วิทัศน์ สำหรับคำนวณหาจุดสนใจ (Key point) ในรูปภาพหนึ่ง ๆ และคำนวณหาคุณลักษณะของ Key point หนึ่งๆ ที่หาได้ SIFT คือการเอาจุดเด่นในรูปที่ไม่ขึ้นอยู่กับการกำหนดทิศทาง ตำแหน่ง มุมการมอง แสงสว่าง เงา ซึ่งจะทำให้สามารถนำมาใช้ในการเปรียบเทียบจุดเด่นในรูปอื่น ๆ ได้ ง่ายและถูกต้องแม่นยำมากยิ่งขึ้นอัลกอริทึมนี้ถูกคิดค้นโดย ศร.เดวิด โวลว์ (David Lowe) แห่ง University British Columbia ในปี 1999 ภาประยุกต์ใช้งาน SIFT มีหลากหลาย ดังเช่น การรู้จำ วัตถุ (Object recognition) การสร้างแผนที่สำหรับนำทางหุ่นยนต์ (Robotic mapping and navigation, image stitching), การสร้างโมเดล 3 มิติของวัตถุหรือทัศนียภาพ (3D modeling of object and scenes) โดยทั่วไปแล้ว Key point จะหมายถึงจุดพิกเซลในภาพที่มีการเปลี่ยนแปลงเชิงสองมิติ (Two -Dimensional) ของระดับความสว่าง (Pixel Intensity) รอบๆ Key point นั้นๆ

2.9 ส่วนประกอบของอัลกอริทึมในการหา SIFT Key point ในภาพหนึ่ง

ส่วนของการตรวจหา Key point ออกมาจากภาพอินพุต (Key point Detection) ซึ่งในส่วนนี้เราจะได้ชุดของพิกัด x, y ของจุดที่จะเป็น Key point ซึ่งจะใช้ในการให้คำอธิบายของ Key point นั้นๆในขั้นตอนถัดไป

ส่วนของการให้คำอธิบายแก่ Key point หนึ่งๆ (Key point Description) ในส่วนนี้ อัลกอริทึมจะคำนวณหาเวกเตอร์อธิบาย (Descriptor Vector) ซึ่งคำนวณมาจากค่าความสว่างของพิกเซลในอาณาบริเวณรอบๆ Keypoint เวกเตอร์อธิบายเหล่านี้จะใช้ในการเปรียบเทียบเพื่อ วัตถุประสงค์ของการรู้จำสัญลักษณ์ (Identity) เมื่อมันไปปรากฏอยู่ในภาพอื่น

2.10 การนำ SIFT ไปประยุกต์ใช้ในการรู้จำวัตถุ

ทำการหา SIFT Key point และ Descriptor Vector จากภาพต้นแบบของวัตถุ (Template Image) ขั้นตอนนี้โดยทั่วไปนิยมเรียกว่า ขั้นตอนการเรียนรู้ (Learning Phase) จาก

ขั้นตอนการทำ SIFT Key point เมื่อเราต้องการจะรู้จำวัตถุที่ปรากฏอยู่ในภาพอินพุตหนึ่ง ๆ ที่ไม่ใช่ต้นแบบ เราสามารถทำได้โดยหา Key point ในภาพอินพุตนั้นๆ ไปเปรียบเทียบกับ Key point หนึ่ง ๆ ใน โมเดลของวัตถุ โดยการเปรียบเทียบจะทำการคำนวณหา Euclidean Distance ระหว่างเวกเตอร์อธิบายที่ประกอบด้วย Key point โดยที่ถ้าค่าระยะทางยิ่งน้อยแสดงว่า Key point ทั้งสองที่นำมาเปรียบเทียบกันยิ่งเหมือนกัน

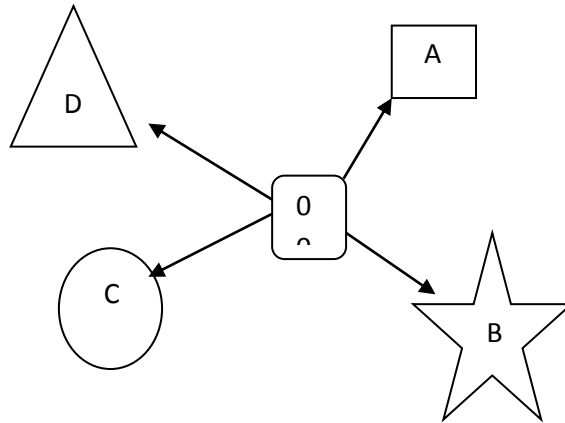
จากชุดของการจับคู่ (Matches) ระหว่าง Key point ในภาพอินพุตกับ Key point ของ โมเดลของวัตถุ เราสามารถประยุกต์ใช้อัลกอริทึมที่ใช้ในการพิสูจน์ว่าเป็นวัตถุที่แท้จริงได้

2.11 Classification

Classifications เป็นขั้นตอนในการจำแนกและตัดสินใจว่าอินพุตที่เข้ามานั้นเป็นภาพรถชนิดใด โดยในขั้นตอนนี้มีหลายวิธีการด้วยกัน เช่น การเปรียบเทียบอินพุตกับโครงสร้างของภาพรถยนต์พื้นฐานข้อมูล การเปรียบเทียบอินพุตกับกฎเพื่อการตัดสินใจ การใช้โครงข่ายประสาทเทียม การเรียนรู้ของเครื่องจักร การคำนวณเชิงวิวัฒนาการ ในส่วนของเนื้อหาที่เกี่ยวข้องกับการทำโปรเจกจะเป็นส่วนของการเรียนรู้ของเครื่องจักร

การเรียนรู้ของเครื่องจักร (Machine Learning) เป็นเทคนิคหรือกระบวนการที่ใช้สำหรับปรับแต่งเครื่องจักรหรือคอมพิวเตอร์ให้มีพฤติกรรมเฉพาะตัวที่สนับสนุนการเรียนรู้ โดยจะมีเทคนิคต่างๆ หลายวิธี แต่ที่เทคนิคที่สนใจคือ เทคนิค Nearest Neighbor Classification

Nearest Neighbor Classification เป็นการเรียนรู้ประเภท Unsupervised Learning ซึ่งเป็นการจำแนกหรือจัดกลุ่มที่มีวิธีการไม่ซับซ้อน โดยพิจารณาจากชุดข้อมูลที่ใกล้เคียงกัน กล่าวคือ ข้อมูลที่ใกล้เคียงกับค่าของข้อมูลที่พิจารณามากที่สุด ในที่นี้ค่าความใกล้เคียงจะหมายถึง ระยะทาง ที่มีค่าน้อยที่สุดระหว่างชุดข้อมูลกับข้อมูลที่พิจารณา



รูปที่ 2.8 แสดงตัวอย่างระยะทางระหว่างชุดข้อมูล

จากรูปที่ 2.8 จะเห็นได้ว่า ข้อมูลที่มีระยะทางใกล้กับจุดศูนย์กลางมากที่สุด คือ A ซึ่งมีระยะเป็น 1 ในทฤษฎีของ Nearest Neighbor การจำแนกชุดข้อมูลที่มีระยะห่าง 1 จะเรียกว่า 1NN (One Nearest Neighbor) โดยระยะห่างของข้อมูลนั้นสามารถกำหนดได้ว่าต้องการมากน้อยเพียงใดด้วยเหตุนี้จึงมีการจำแนกที่เรียกว่า k NN ซึ่ง k แทนด้วยค่าระยะทางระหว่างข้อมูลที่ต้องการ สำหรับการหาค่าระยะทางจะใช้สมการทางทฤษฎีการวัดค่าของ Euclidean คือ

$$d = \sqrt{\sum_{i=1}^n (p_i - q_i)^2}$$

สมการทางทฤษฎีการวัดค่าของ Euclidean

2.12 OpenCV (Open Source Computer Vision Library)

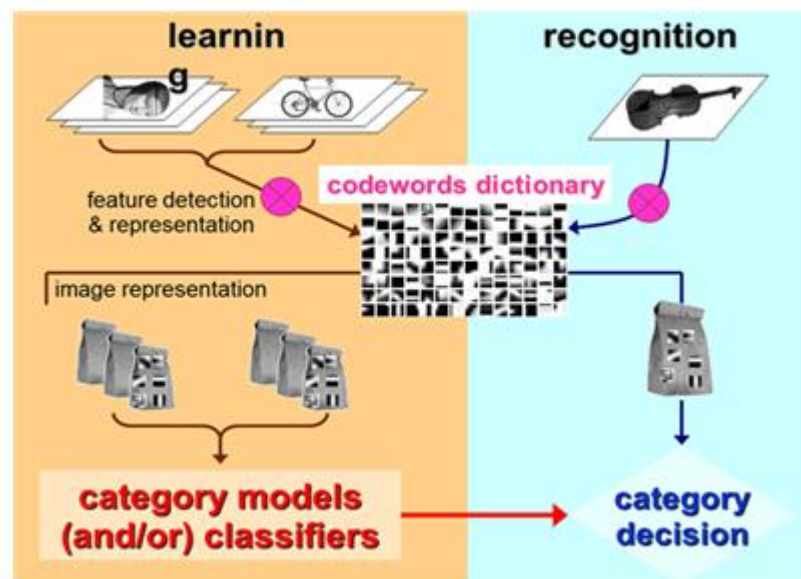
OpenCV หรือ Open Source Computer Vision Library ใช้ในการประมวลผลภาพและงานทางด้าน การมองเห็นของคอมพิวเตอร์ (Computer Vision) Library นี้ถูกพัฒนาขึ้นด้วยภาษา C และ C++ และยังมี interface ที่ไว้เชื่อมต่อกับ tool อื่นด้วยเช่น Python, Ruby, Matlab เป็นต้น นอกจากนี้ OpenCV เป็น library ที่สร้างขึ้นเพื่อให้ผู้ใช้หรือนักพัฒนาสามารถใช้ฟังก์ชันใน library มาพัฒนาชิ้นงานที่มีความซับซ้อนโดยใช้เวลาเพียงไม่นาน OpenCV ประกอบด้วย Data Structure และ Algorithm

- Data Structure ใช้เก็บข้อมูลต่าง ๆ อาทิ เช่น รูปภาพ เมทริกซ์ พิกัด

- Algorithm เพื่อการประมวลผลต่าง โดยเฉพาะการประมวลผลทางรูปภาพ

ข้อจำกัดของ OpenCV คือ สามารถใช้งานได้เฉพาะเครื่องคอมพิวเตอร์ที่ใช้หน่วยประมวลผล (CPU) จาก Intel Corporation แต่ข้อจำกัดนี้ทำให้เกิดจุดเด่นเช่นกัน กล่าวคือ การประมวลผลต่าง ๆ จะใช้ความสามารถของหน่วยประมวลผลอย่างเต็มที่ ทำให้โปรแกรมที่พัฒนาโดยใช้ OpenCV นี้มีประสิทธิภาพในการประมวลผลที่สูงมาก

2.13 หลักการทำงานโครงสร้างแบบถุงคำ (bag of words)



รูปที่ 2.9 ส่วนของโครงสร้างแบบถุงคำ

ส่วนที่หนึ่ง Image representation หลักการทำงานในรูป (ด้านซ้าย) แสดงให้เห็นว่าเมื่อเรารับภาพเข้ามา ภาพที่เรารับเข้ามานั้นก็จะถูกนำไปตรวจสอบในฐานข้อมูล ว่าเป็นภาพที่มีส่วนของภาพ คล้ายกับภาพที่เรารับเข้ามาหรือไม่ถ้ามีก็จะแสดงออกมาว่า ภาพที่เรารับเข้ามานั้นเป็นภาพอะไร

ส่วนที่สอง ส่วนของการรู้จำ (Recognition) จะเป็นส่วนที่ทำการค้นหาว่า input ที่รับเข้ามา มีอยู่ในฐานข้อมูลหรือไม่ โดยสามารถอธิบายด้วยรูปภาพได้ดังรูป (ด้านขวา) ส่วนของ

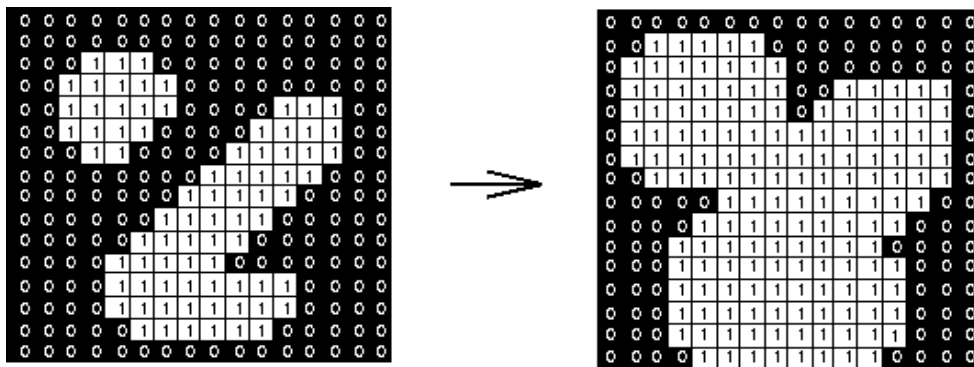
ระบบรู้จำจะเป็นส่วนที่นำข้อมูลมาทำการ segmentation แล้วนำไป เลือกเฉพาะส่วนที่เป็น จุดเด่นของภาพนำไปเก็บลงในฐานข้อมูลโดยจะถูกเรียกใช้ก็ต่อเมื่อ มีการรับ input เข้ามา

2.14 การเปลี่ยนแปลงลักษณะรูปร่างหรือโครงร่างของภาพ (Morphological processing)

การเปลี่ยนแปลงลักษณะรูปร่างหรือโครงร่างของภาพ คือการประมวลผลของภาพ ทางด้านโครงสร้าง โดยเกี่ยวกับการแยกส่วนประกอบของภาพออกเพื่อใช้ในการแสดงรูปร่าง ใน เมทริกซ์จะประกอบไปด้วยค่าระดับขาว-ดำ 2 ค่า คือ 0 และ 1

2.14.1 การขยาย (Dilation)

การขยาย (Dilation) จะพิจารณาข้อมูลภาพซึ่งเป็นภาพขาว-ดำ เป็นการขยายภาพให้ ใหญ่ขึ้น เพื่อเพิ่มสีให้กับวัตถุที่แสดงผลในขั้นตอนสุดท้าย ซึ่งการขยายวัตถุจะทำได้โดยการ กำหนดส่วนประกอบโครงสร้าง (Structuring element) และนำส่วนประกอบโครงสร้างไปกราด บนข้อมูลภาพตามลำดับตลอดทั้งภาพ โดยเมื่อจุดเริ่มต้นของส่วนประกอบโครงสร้างหรือจุด กำเนิดตรงกับตำแหน่งข้อมูลภาพที่เท่ากับ 1 จะทำการยูเนียนส่วนประกอบโครงสร้าง เข้ากับ ข้อมูลภาพดังรูป

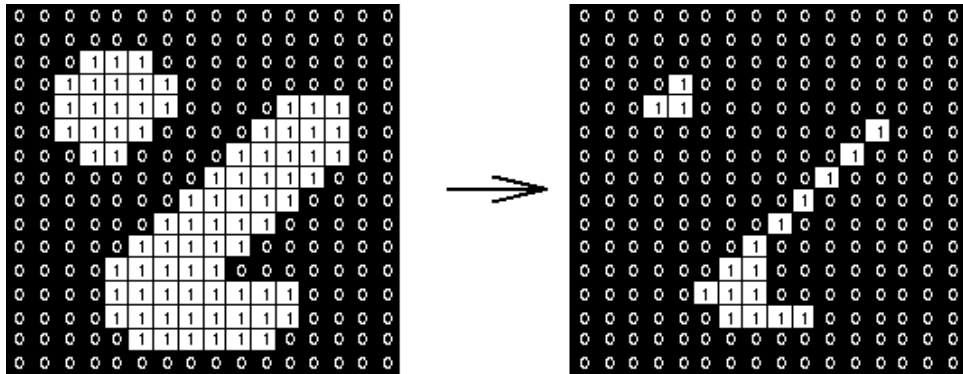


รูปที่ 2.10 ส่วนของการขยาย

2.14.2 การกร่อนขนาด (Erosion)

การกร่อนขนาด (Erosion) เป็นการกร่อนขนาดบริเวณขอบของวัตถุ ซึ่งการกร่อนมีวิธี คล้ายกับการขยายคือ สร้างส่วนประกอบโครงร่างขึ้นมาแล้วนำไปกราดตามข้อมูลภาพ โดยจะ เลื่อนไปทุกตำแหน่งเปรียบเทียบกับข้อมูลภาพ ถ้าข้อมูลมีค่าเหมือนกับส่วนประกอบโครงร่างจะ

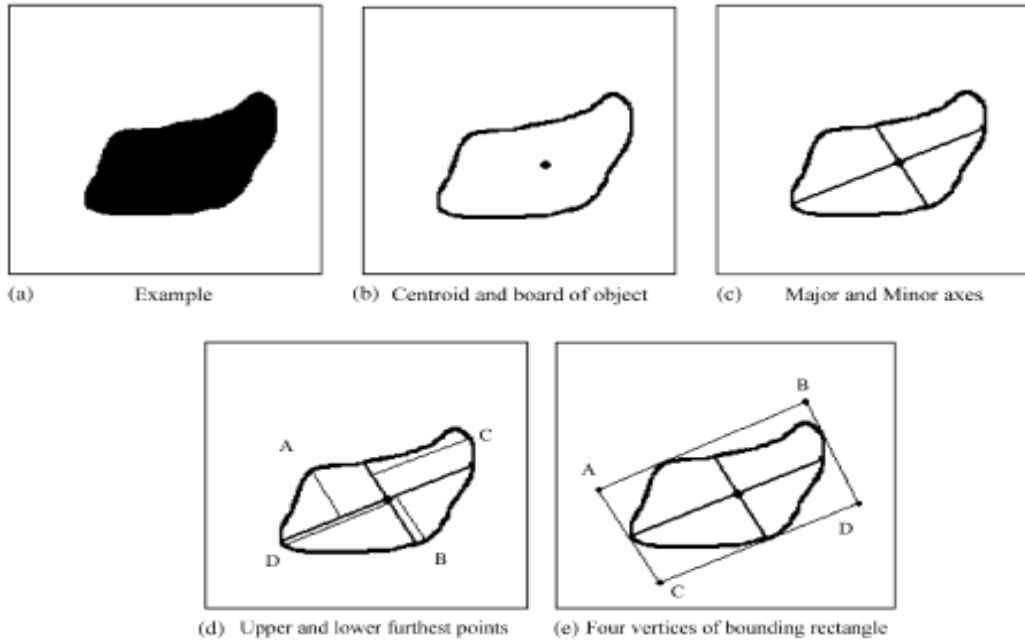
ทำการกำหนดค่าข้อมูลภาพที่ตรงกับตำแหน่งที่ตรงกับจุดเริ่มต้นหรือ จุดกำเนิดของส่วนประกอบ
โครงร่างให้เท่ากับ 1



รูปที่ 2.11 ส่วนของการกร่อนขนาด

2.15 สีเหลี่ยมในการปิดล้อมวัตถุ (Bounding Box)

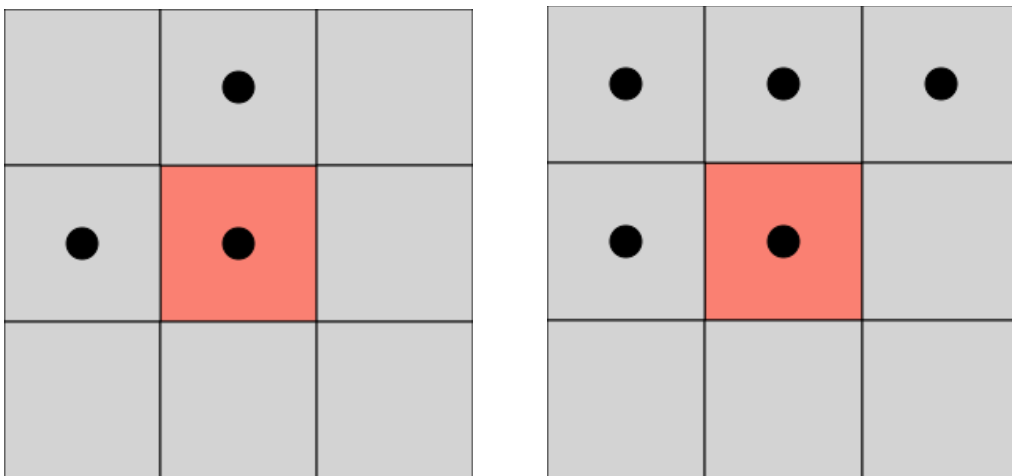
คือกล่องสี่เหลี่ยมสำหรับปิดล้อมรอบบริเวณนั้นๆ หรือบริเวณที่ถูกกำหนดหมายเลข ซึ่งกล่องสี่เหลี่ยมนี้จะช่วยในการคำนวณจุดศูนย์กลาง พื้นที่ ความยาวแกนเอก-แกนโท จุดพิกัด ฯ ของบริเวณหรือวัตถุในภาพที่สนใจ โดยจากรูปที่ 2.12 จะแสดงแนวความคิดการสร้างกรอบภาพ โดยรูป 2.12(a) คือภาพวัตถุสีขาว-ดำ ส่วนในรูปที่ 2.12(b) คือขอบและจุดศูนย์กลาง (Centroid) ของวัตถุ ต่อจากนั้นจะมีการหาแกนเอกและแกนโทของวัตถุดังรูปที่ 2.12(c) จากนั้นจะมีการหาจุดที่มีระยะห่างจากจุดศูนย์กลางวัตถุถึงขอบวัตถุที่ยาวที่สุดโดยพิจารณาจากส่วนด้านบนของแกนเอกและส่วนล่างของแกนเอก (A และ B) รวมทั้งการหาจุดที่มีระยะห่างจากจุดศูนย์กลางวัตถุถึงขอบวัตถุที่ยาวที่สุดโดยพิจารณาจากบริเวณส่วนซ้ายของแกนโทและส่วนขวาของแกนโท (C และ D) สุดท้ายก็จะได้จุดพิกัดในการสร้างกล่องสี่เหลี่ยมในการปิดล้อมวัตถุ (Minimum bounding box) นั้น



รูปที่ 2.12 แนวความคิดการสร้างสี่เหลี่ยมในการปิดล้อมวัตถุ

2.16 การกำหนดหมายเลขให้ส่วนที่เชื่อมกัน (Connected-component labeling)

ใช้ในการวิเคราะห์ส่วนประกอบหรือบริเวณที่มีการเชื่อมติดกันและใช้ในการแยกบริเวณต่างๆออกจากกันซึ่งไม่เหมือนกับวิธีตัดแยกโดยวิธีการกำหนดหมายเลขให้ส่วนที่เชื่อมกัน (Connected-component labeling) เป็นวิธีที่ใช้ในคอมพิวเตอร์วิทัศน์ (Computer vision) เพื่อตรวจจับบริเวณที่เชื่อมติดกันของภาพขาวดำ ภาพสี แต่โดยทั่วไปใช้ในภาพขาวดำที่ผ่านการประมวลผลมาแล้ว อาจใช้ในการนับจำนวน การกรองและการติดตาม



รูปที่ 2.13 แบบจุด 4 จุด เชื่อมกันและแบบจุด 8 จุด เชื่อมกัน

จากรูปที่ 2.14 จุดที่อยู่รอบๆ จุดที่อยู่ตำแหน่งตรงกลางคือ เพื่อนบ้านมีทั้งแบบจุด 4 จุด เชื่อมกัน และแบบจุด 8 จุด เชื่อมกันโดยเพื่อนบ้านที่เชื่อมกันก็คือขอบนั่นเอง การกำหนดหมายเลขมี 2 แบบคือแบบจุด 4 จุด เชื่อมกันและแบบจุด 8 จุด เชื่อมกัน มีข้อแตกต่างกันคือ ถ้าเป็นแบบจุด 4 จุด เชื่อมกันจะเลือกจุดที่เชื่อมกันเฉพาะบน ล่าง ซ้าย ขวา แต่หากเป็น แบบจุด 8 จุด เชื่อมกันจะเลือกรวมจุดที่อยู่มุมทแยงจากตำแหน่งกลางอีก 4 จุดด้วย

	u	
l	p	

รูปที่ 2.14 อักษรตำแหน่งจุดภาพ

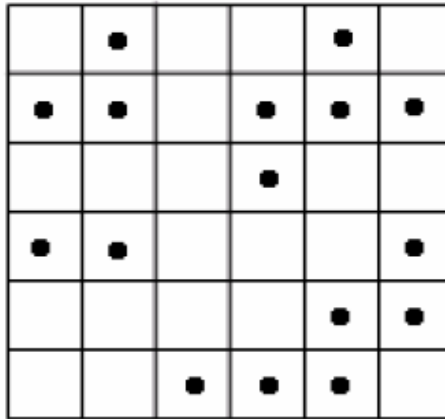
ดังรูปที่ 2.14 กำหนดให้ p (pixel) แทนจุดภาพที่กำลังพิจารณา u แทนจุดภาพที่อยู่ตำแหน่งเหนือจุด p l แทนจุดภาพที่อยู่ตำแหน่งซ้ายของจุด p จะเริ่มทาจากซ้ายไปขวาและบนลงล่าง จากนั้นทำการกำหนดเลขหมายตามขั้นตอนดังนี้

ขั้นที่ 1. จากรูปที่ 2.14 ถ้าจุด p ไม่ใช่จุดภาพให้เลื่อนจุดในตำแหน่งถัดไป ถ้า p เป็นจุดภาพให้ตรวจสอบสถานะของ u และ l ถ้าไม่มีจุดใดเป็นจุดภาพให้ กำหนดหมายเลขใหม่ให้กับจุด p ถ้ามี 1 จุดเป็นจุดภาพให้นำหมายเลขของ จุดนั้นมากำหนดให้จุด p แต่ถ้ามีมากกว่า 1 จุดเป็นจุดภาพสามารถหาหมายเลขของจุดใดก็ได้มากำหนดให้จุด p โดยถือว่าทุกหมายเลขเทียบเท่ากัน

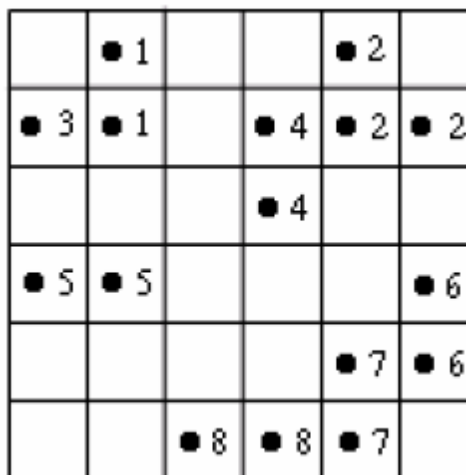
ขั้นที่ 2. เมื่อสิ้นสุดการสแกนในขั้นที่ 1 จุดภาพทุก ๆ จุดจะมีหมายเลขกำหนดดังรูปที่ 2.16 แต่บางหมายเลขจะเทียบเท่ากัน ให้รวมกลุ่มหมายเลขที่เทียบเท่ากันดังรูปที่ 2.17 จากนั้น กำหนดหมายเลขให้แต่ละกลุ่ม

ขั้นที่ 3. นำหมายเลขของแต่ละกลุ่มจากขั้นที่ 2 ไปแทนหมายเลขของจุดภาพที่อยู่ในกลุ่มเดียวกัน

จากผลลัพธ์ที่ได้จากรูปที่ 2.15 จะเห็นว่าในแต่ละบริเวณจะมีหมายเลขกำกับไม่ซ้ำกัน
 ดังนั้นหมายเลขสูงสุดก็คือจำนวนบริเวณทั้งหมดที่เชื่อมกันนั่นเอง



รูปที่ 2.15 ตัวอย่างจุดภาพและตำแหน่ง



รูปที่ 2.16 หมายเลขของแต่ละจุดภาพตามขั้นตอนที่ 1

Set ID	Equivalent Labels
1	1,3
2	2,4
3	5
4	6,7,8

รูปที่ 2.17 กลุ่มร่วมที่หมายเลขเทียบเท่ากัน

	● 1			● 2	
● 1	● 1		● 2	● 2	● 2
			● 2		
● 3	● 3				● 4
				● 4	● 4
		● 4	● 4	● 4	

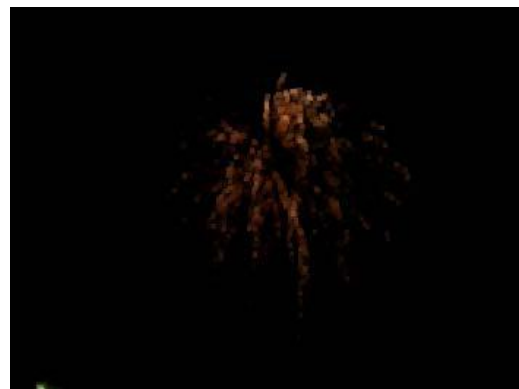
รูปที่ 2.18 หมายเลขของแต่ละจุดภาพตามชั้นที่ 3

2.17 Opening and Closing

2.17.1 Opening ใช้เพื่อกำจัดรายละเอียดขนาดเล็กของภาพ และการทำ Opening จะทำให้พิทเชลของภาพจะถูกเปิดกว้างมากขึ้นดังภาพและวิธีการของ Opening คือการทำ Erosion ก่อน จากนั้นจึงทำ Dilation ดังสมการ $A \circ B = (A \ominus B) \oplus B$ เรียก B ว่าเป็น Structuring element



(ก)



(ข)

รูปที่ 2.19 ภาพการทำ Opening (ก) ภาพต้นฉบับ (ข) ผลลัพธ์จากการทำ Opening

2.17.2 Closing ทำในวิธีตรงข้ามกับ Opening จะเป็นการทำให้ภาพมีการเชื่อมต่อกันมากขึ้นและการทำ Closing จะทำให้พิกเซลของภาพจะถูกปิดเชื่อมต่อกันมากขึ้นดังภาพที่ 2-15

วิธีการทำ Closing คือการทำ Dilation ก่อนจึงทำ Erosion ดังสมการ $A \bullet B = (A \oplus B) \ominus B$ เรียก B ว่าเป็น Structuring element



(ก)



(ข)

รูปที่ 2.20 ภาพการทำ Closing (ก) ภาพต้นฉบับ (ข) ผลลัพธ์จากการทำ Closing