

บทที่ 2

ทฤษฎีที่เกี่ยวข้อง

2.1 คอมพิวเตอร์วิทัศน์ (computer vision)

คอมพิวเตอร์วิทัศน์ (computer vision) เป็นสาขาหนึ่งของปัญญาประดิษฐ์จุดประสงค์หลักของคอมพิวเตอร์วิทัศน์ คือการทำให้คอมพิวเตอร์สามารถ "เข้าใจ" ทิวทัศน์ หรือคุณลักษณะต่างๆ ในภาพได้

เป้าหมายของคอมพิวเตอร์วิทัศน์ โดยทั่วไปคือ การทำให้คอมพิวเตอร์สามารถเข้าใจภาพที่รับเข้ามาโดยการแปลงจากภาพ (image) เป็นแบบจำลอง (model) ที่คอมพิวเตอร์สามารถรับรู้คุณสมบัติที่ต้องการได้ เช่น คุณสมบัติทางรูปร่าง เพื่อการจดจำ เพื่อการหยิบจับ เป็นต้น

กระบวนการจะเกี่ยวเนื่องตั้งแต่ การรับภาพ (ซึ่งไม่ได้จำกัดอยู่เพียงแต่ภาพที่รับด้วยกล้องแต่อาจหมายถึงภาพถ่ายความร้อน และ ภาพจากอุปกรณ์วัดระยะ ฯลฯ) การประมวลผลภาพ การสร้างแบบจำลองจากภาพ และการรับรู้เข้าใจของคอมพิวเตอร์จากแบบจำลอง

งานที่พบเห็นทั่วไปของ คอมพิวเตอร์วิทัศน์คือ การตรวจจับ แยกแยะ หาตำแหน่ง และจดจำ วัตถุที่ต้องการในภาพ เช่น การจดจำหน้าคน การสร้างแบบจำลองของวัตถุ เพื่อจุดประสงค์ต่างๆ เช่น การใช้แขนกลหยิบจับ ตัด หรือ เพื่อการตรวจสอบ เช่น การสร้างภาพสามมิติทางการแพทย์ การสร้างแบบจำลองสามมิติของสภาพแวดล้อม เช่น เพื่อการรับรู้ตำแหน่ง และการเคลื่อนที่ของหุ่นยนต์

งานของคอมพิวเตอร์วิทัศน์ จะต้องใช้ความรู้จากสาขาต่างๆ คือ การรู้จำแบบ การเรียนรู้เชิงสถิติ projective geometry การประมวลผลภาพ ทฤษฎีกราฟ และสาขาอื่นๆ

มีความพยายามในการสร้างระบบที่มีความสามารถในการวิเคราะห์และประมวลผลภาพ โดยเป้าหมายหลักประการหนึ่งก็คือการสร้างระบบที่สามารถบอกได้ว่าภาพสองมิติที่ได้รับมานั้นเป็นภาพของวัตถุใด มีลักษณะในสามมิติอย่างไร ศาสตร์ที่ว่ามีชื่อเรียกกันว่า computervision แต่ computervision ยังไม่สามารถพัฒนาให้ดีพอได้เพราะธรรมชาติของปัญหาเองที่ต้องการกู้คืน ข้อมูลสามมิติจากข้อมูลเข้าสองมิติ ซึ่งเป็นไปไม่ได้ นอกจากมีข้อกำหนดเสริมที่เพียงพอ

2.2 อุปกรณ์รับภาพ

ภาพที่สามารถนำมาวิเคราะห์หาข้อมูลต้องมีวิธีการบันทึกภาพที่เหมาะสม โดยหลักการคร่าวๆของการบันทึกภาพก็เหมือนกับการทำงานของกล้องถ่ายรูป คือแสงจากต้นทางที่ตกกระทบวัตถุจะสะท้อนและผ่านจตุรวมแสง โดยแสงที่ผ่านจตุรวมแสงจะไปตกกระทบที่ส่วนรับภาพ จากนั้นภาพที่ได้จะถูกนำไปประมวลผลต่อไป

2.2.1 อุปกรณ์รับภาพธรรมชาติ มีดังนี้คือ

2.2.1.1 ตา

เป็นอวัยวะของมนุษย์ที่ทำหน้าที่รับภาพโดยมีส่วนประกอบดังนี้ คือ Extraocular Muscles เป็นกล้ามเนื้อทำหน้าที่ควบคุมการเคลื่อนไหวของลูกตา เพื่อให้อยู่ในทิศทางที่จะรับภาพที่ต้องการ ได้เสมือนการปรับมุมกล้องให้ได้พอเหมาะ Cornea ทำหน้าที่เป็นส่วนควบคุมปริมาณแสงที่เข้าสู่ตาให้เหมาะสมเสมือนกับการปรับรูรับแสงของกล้องถ่ายรูปที่จะให้แสงเข้า Lens ทำหน้าที่เป็นส่วนรวมแสงเหมือนเลนส์กล้อง Retina เป็นฉากรับแสง ประกอบไปด้วยเซลล์ไวแสงแบบต่างๆเหมือนฟิล์มในกล้องถ่ายรูป Macula ทำหน้าที่ปรับให้เกิดความคมชัดของภาพที่ retina มาเสมือนการปรับ f-stop ของกล้องถ่ายรูป Optic nerve ทำหน้าที่ส่งต่อข้อมูลที่รับมาสู่สมอง

2.2.1.2 กล้องรูเข็ม

เป็นอุปกรณ์รับภาพที่อาศัยรูเล็กๆ เพียงรูเดียวให้แสงสามารถลอดผ่านได้ โดยแสงที่กระทบวัตถุจะลอดผ่านรูนี้เป็นเส้นตรง และไปกระทบกับฉากทำให้เกิดเป็นภาพของวัตถุข้อเสียของอุปกรณ์ชิ้นนี้คือ ความเข้มแสงที่ตกกระทบวัตถุมีผลต่อความชัดเจนของภาพที่ได้เป็น อย่างมาก ขนาดของรูของกล้องที่ไม่เหมาะสมก็เป็นปัจจัยที่ทำให้ภาพพร่ามัวได้

2.2.1.3 เลนส์

ปริมาณแสงที่ผ่านรูของกล้องรูเข็ม ไปยังฉากมีจำกัด เพราะการขยายขนาดรูให้ใหญ่เพื่อให้ได้ปริมาณแสงที่เพิ่มขึ้นอาจทำให้ได้ภาพมัว ปริมาณแสงที่จำกัดนี้เป็นสาเหตุให้ได้ภาพที่มีมืดและไม่ชัดเจน เราสามารถแก้ปัญหานี้ได้ด้วยการใช้เลนส์เป็นอุปกรณ์รวมแสง โดยใช้หลักการหักเหของแสง (refraction) แทนการให้

แสงรอดผ่านรูการทำงานของเลนส์ จะกล่าวถึงการหักเหของแสง เมื่อเดินทางผ่านตัวกลางที่มีความทึบแสงต่างกัน สำหรับแท่งแก้วใสที่มีผิวเป็นส่วนโค้งกลม (spherical surface) ถ้าแสงขนานที่เดินทางในอากาศและตกกระทบ แท่งแก้วแล้วหักเหและมาพบกันที่จุดโฟกัส เราเรียกระยะห่างที่ใกล้ที่สุดจากจุดนี้ไปยังผิวของแก้วว่า ระยะโฟกัส

2.2.1.4 กล้อง CCD

กล้องถ่ายรูปบันทึกภาพลงบนฟิล์ม โดยอาศัยปฏิกิริยาเคมี ที่ทำให้เกิดการเปลี่ยนแปลงบนส่วนของฟิล์มที่ได้รับแสง แต่ภาพที่บันทึกบนแผ่นฟิล์ม นั้นไม่สามารถที่นำไปประมวลผลด้วยคอมพิวเตอร์โดยตรงได้ กล้อง CCD (charged coupled device) ประกอบด้วยส่วนรับแสง และส่วนบันทึกภาพซึ่งเป็นวงจรรวมที่ประกอบไปด้วยเซลล์เล็กๆ เรียงอยู่อย่างมีระเบียบ โดยแต่ละเซลล์จะมีความไวแสงและเปลี่ยนแสงที่มากระทบให้อยู่ในรูปของสัญญาณไฟฟ้าพร้อมที่จะส่งไปยังคอมพิวเตอร์เพื่อประมวลผล

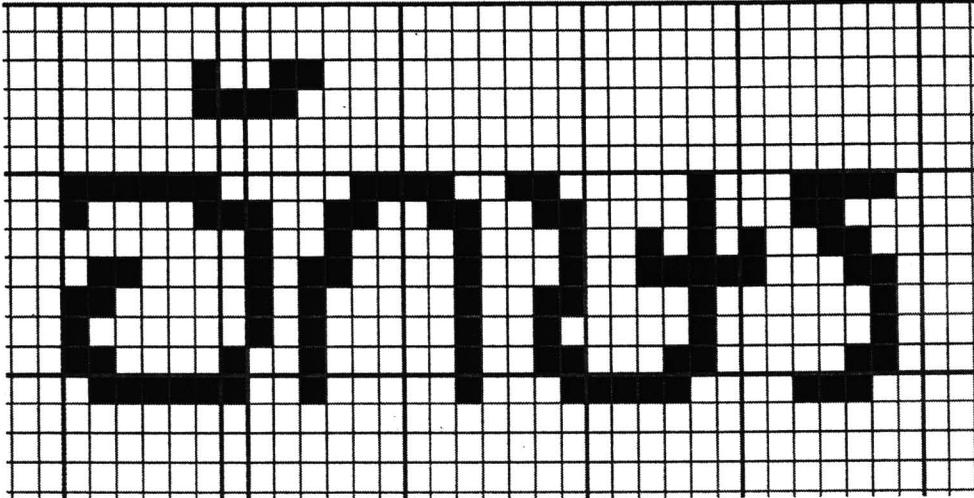
2.3 การวิเคราะห์และประมวลผลภาพ

การวิเคราะห์ประมวลผลภาพและการรับรู้จากภาพ เป็นศาสตร์ที่ครอบคลุมเนื้อหา กว้างขวางและลึกซึ้งอย่างมาก เพื่อให้เกิดความเข้าใจถึงลักษณะของปัญหาที่สามารถประยุกต์ใช้เทคนิคต่างๆ จากศาสตร์แขนงนี้ จะขอยกตัวอย่างการประมวลผลภาพแบบพื้นฐาน โดยจะขอกกล่าวถึงการประมวลผลภาพไบนารี (binary image) ซึ่งเข้าใจได้ง่ายก่อน แล้วจึงจะกล่าวถึงการประมวลผลภาพเกรย์สเกล (gray scale image) ที่มีความซับซ้อนขึ้นเป็นลำดับต่อไป

2.3.1 ภาพไบนารี

ภาพไบนารีคือ ภาพซึ่งในแต่ละพิกเซลจะมีความเข้มของแสงได้สองแบบเท่านั้น คือสว่างกับมืด โดยอาจเขียนแทนได้ด้วยเลข 0 และ 1 ด้วยความเข้มของแสงเพียงสองระดับ จึงทำให้มีข้อจำกัดที่จะนำมาใช้แสดงภาพต่างๆ ไปแต่อย่างไรก็ตามการมีความเข้มของแสงเพียงสองระดับก็ทำให้การประมวลผลทำได้ง่ายมีประสิทธิภาพ การประมวลผลภาพไบนารีนำไปใช้มากในการประมวลผลเอกสาร การประมวลผลภาพในอุตสาหกรรมที่ต้องการความเร็วสูง การมีความเข้มเพียงสองระดับในภาพไบนารี ทำให้สามารถที่จะเลือกพิจารณาให้ความเข้มระดับหนึ่งแทนภาพของสิ่งที่เราสนใจโดยเรียกว่า “พิกเซลภาพ” และความเข้มอีกระดับแทนพื้นหลังจะถูกรู้เรียกว่า “พิกเซลพื้นหลัง”

ในทางปฏิบัติการพิจารณาภาพไบนารี เราจะไม่สนใจแต่ละพิกเซลแยกกันไป แต่จะสนใจกลุ่มของพิกเซลที่อยู่ติดกัน หรือที่เรียกว่า “พิกเซลเพื่อนบ้าน” เช่นกลุ่มของพิกเซลที่เรียงกันเป็นตัวอักษร ดังรูปที่ 2.1



รูปที่ 2.1 ภาพกลุ่มพิกเซลภาพที่ประกอบเป็นตัวอักษร

3	2	1
4	P	0
5	6	7

รูปที่ 2.2 พิกเซล p และพิกเซลรอบๆ

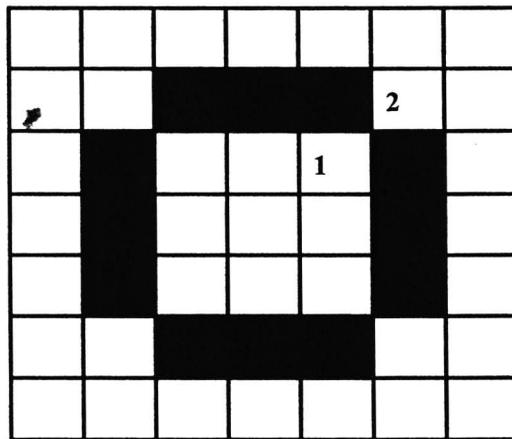
2.3.2 สำหรับการตีความว่าพิกเซลใดอยู่ติดกันทำได้สองแบบหลักคือ

2.3.2.1 แบบสี่เพื่อนบ้าน (4-connectedness)

จะถือว่าแต่ละพิกเซลอยู่ติดกับสี่พิกเซลที่อยู่ด้านบน, ล่าง, ซ้ายและขวา ในบางครั้งอาจใช้ชื่อทิศในการเรียกพิกเซลรอบๆ เช่นเรียกพิกเซล 2 ว่าพิกเซลเหนือของพิกเซล p เป็นต้น

2.3.2.2 แบบแปดเพื่อนบ้าน (8-connectedness)

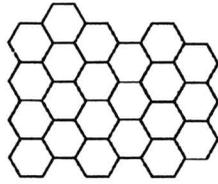
แต่ละพิกเซล ที่อยู่ติดกับทั้งแปดพิกเซลที่อยู่ล้อมรอบนิยามของการอยู่ติดกันของพิกเซลทั้งสองแบบที่กล่าวมานิยมใช้การประมวลผลภาพไบนารีแต่มีสิ่งที่จะต้องคำนึงคือความไม่สอดคล้องบางประการกับเรขาคณิต คือ ทฤษฎีเส้นโค้งของจอร์แดน (Jordan's curve theorem) เส้นโค้งบนระนาบที่ลากเป็นวงปิดจะแบ่งระนาบออกเป็นสองส่วนที่ไม่ติดกัน คือส่วนที่อยู่ภายในวงและส่วนที่อยู่นอกวง แต่สำหรับภาพไบนารีหากลองกำหนดให้ใช้นิยามการติดกันแบบแปดเพื่อนบ้าน แล้วลองลากเส้นให้เป็นวงเช่นในรูปที่ 2.3 จะเห็นได้ว่าพิกเซลที่ประกอบเป็นเส้นที่ลากนี้ไม่ได้แบ่งพิกเซลที่เหลือให้เป็นสองส่วนที่ไม่ติดกันเลย สังเกตพิกเซล 1 กับ 2 ของพื้นที่หลังที่ถือว่าติดกัน ในลักษณะเดียวกันหากเราใช้การติดกันแบบสี่เพื่อนบ้านกับภาพไบนารีในรูปที่ 2.3 ภาพนี้จะประกอบไปด้วยเส้นสี่เส้นที่ไม่ติดกันและไม่มีวงปิดเกิดขึ้นแต่กลับทำให้พิกเซลที่เหลือ (พิกเซลพื้นหลัง) ยกกันเป็นสองกลุ่มที่ไม่ติดกัน



รูปที่ 2.3 ภาพไบนารีของเส้นที่ลากเป็นวง

หากความไม่สอดคล้องนี้ไม่ทำให้เกิดปัญหากับงานที่ทำก็ไม่เป็นไร แต่หากจำเป็นต้องแก้ไขให้สอดคล้อง วิธีแก้ปัญหาวีธีหนึ่งก็คือการเปลี่ยนไปใช้พิกเซลที่เป็นรูปหกเหลี่ยมแทนพิกเซลสี่เหลี่ยม (รูปที่ 2.4) และกำหนดให้พิกเซลที่มีขอบร่วมกันเป็นพิกเซลที่ติดกัน อีกวิธีที่ทำได้โดยยังคงรูปแบบพิกเซลสี่เหลี่ยมก็

คือการพิจารณาการติดกันของพิกเซลพื้นหลังแบบสี่เพื่อนบ้าน แต่ใช้การติดกันแบบแปดเพื่อนบ้านสำหรับพิกเซลที่แทนภาพของสิ่งที่สนใจ



รูปที่ 2.4 พิกเซลแบบหกเหลี่ยม

2.4 แนะนำภาพดิจิทัล

รูปภาพที่เห็นกันอยู่ ไม่ว่าจะเป็นภาพที่ถ่ายโดยใช้กล้องธรรมดา หรือแบบดิจิทัล ถ้ามองกันในแบบของคอมพิวเตอร์แล้วละก็ มันก็คือ จุดสีหลาย ๆ จุดที่นำมาเรียงต่อ ๆ กัน จนสามารถบอกได้ว่าเรียงกันเป็นรูปอะไร เนื้อหาของรูปภาพเป็นอย่างไร การมองเพื่อทำความเข้าใจรูปภาพหนึ่ง ๆ ไม่ว่าจะเป็นภาพถ่าย หรือภาพที่เป็นแบบดิจิทัลในคอมพิวเตอร์ก็ตาม ในมุมมองของมนุษย์กับรูปภาพ หรือมุมมองของคอมพิวเตอร์กับรูปภาพ เป็นคนละมุมมองกัน และแตกต่างกันอย่างสิ้นเชิง มนุษย์สามารถเข้าใจถึงเนื้อหาของภาพได้ว่าภาพที่ปรากฏนั้นให้ความพึงพอใจ ความน่าสนใจมากน้อยแค่ไหน และภาพนี้บอกอะไร สามารถสื่อถึงความรู้สึกอะไรบางอย่างได้หรือไม่ และอีกหลายความรู้สึกที่ได้จากการมองภาพ

2.4.1 การแปลงภาพให้เป็นภาพเชิงดิจิทัล

ภาพเป็นกระบวนการทางแสง (Optical Process) ซึ่งเกิดจากพลังงานคลื่นแม่เหล็กไฟฟ้า (Electromagnetic Spectrum) หลายๆ ช่วงความถี่ เช่น แสงธรรมดา รังสีเอ็กซ์เรย์ (X-Ray) รังสีอินฟราเรด (Infrared) เป็นต้น และพลังงานเสียง เช่น อัลตราซาวนด์ (Ultrasound) ตกกระทบวัตถุแล้วสะท้อนกลับมาสู่ประสาทรับรู้ทางตาของมนุษย์ หรืออุปกรณ์ตรวจจับ เช่น เซนเซอร์ (Sensor)

ถ้าพูดกันในภาษาทางเทคนิคแล้ว ภาพดิจิทัลนั้นก็คือ ฟังก์ชัน 2 มิติ หรือ $f(x,y)$ ของค่าความเข้มของแสงโดยที่ x และ y คือ ค่าที่บอกถึงตำแหน่งในระบบพิกัดฉาก และค่าของฟังก์ชัน ณ ตำแหน่งใดๆ จะเป็นสัดส่วนกับความสว่างของแสง ณ ตำแหน่งนั้น กระบวนการแปลงภาพให้เป็นภาพในเชิงดิจิทัลซึ่งเรียกว่า Image Digitization มีกระบวนการ 3 ขั้นตอน คือ การบันทึกภาพ (Image Acquisition), การสุ่มเลือกจุดตำแหน่ง (Image Sampling) และการประมาณค่าความเข้มของแสง (Image Quantization) แต่เมื่อรูปภาพ ถูกนำมาทำเป็นภาพในคอมพิวเตอร์ มันจะรู้และเข้าใจภาพเป็นเพียงแค่เป็นจุดสีหลาย ๆ จุดที่เรียงต่อกันในความสัมพันธ์ระหว่างจุดภาพที่เหมาะสม ภาพดิจิทัลถึงแม้จะ

เก็บอยู่ในรูปของไฟล์ในดิสก์ของคอมพิวเตอร์เอง หรือแม้แต่ว่าเราจะนำเอาภาพสวยๆ มาเป็นวอลเปเปอร์พื้นหลังของ Desktop ใน Windows ก็ไม่อาจจะรู้และเข้าใจถึงเนื้อหาของภาพที่ปรากฏได้ ยกเว้นผู้ใช้คอมพิวเตอร์นั้นจะบอกได้ว่านี่คือภาพที่สวยหรือไม่

2.4.2 การบันทึกภาพ (ImageAcquisition)

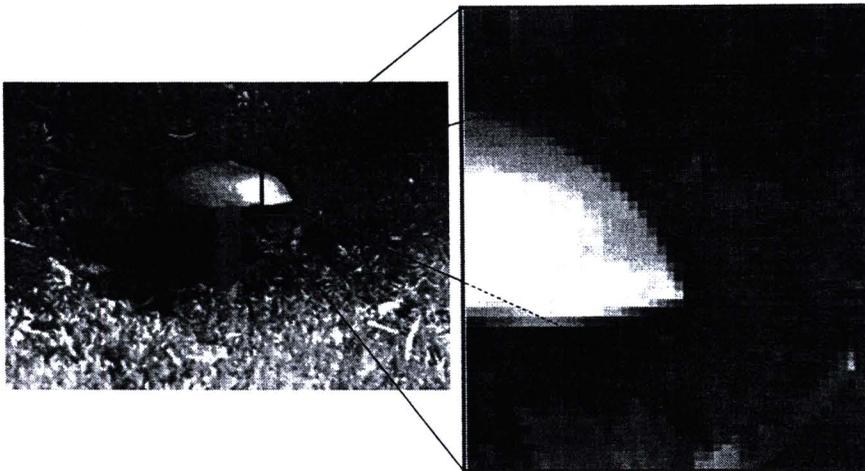
ก่อนที่จะได้ภาพมาอันดับแรก จะต้องทำการถ่ายภาพเสียก่อน การถ่ายภาพ เป็นการแปลงภาพเชิงต่อเนื่อง (ContinuousImage) แบบ 3 มิติ ให้เป็นภาพเชิงต่อเนื่อง 2 มิติ โดยใช้อุปกรณ์เชิงแสง (OpticalDevice) เช่น กล้องถ่ายรูป เพื่อแปลงภาพให้มาเป็นภาพบนฟิล์ม, รูปถ่ายบนกระดาษ หรือภาพบนจอคอมพิวเตอร์ โดยปกติแล้ว ภาพที่ตามองเห็นกันอยู่นั้น มีทั้งความกว้าง, ความสูง และความลึก ซึ่งเป็นแบบ 3 มิตินั่นเอง การถ่ายภาพด้วยกล้องจะทำให้ได้ภาพมา แต่จะเป็นภาพที่มีแต่ความกว้าง และความสูงเท่านั้น ซึ่งเป็นภาพแบบ 2 มิติ เพราะไม่อาจจะถ่ายความลึกของสถานที่มาได้ ในส่วนของการบันทึกภาพนี้ เป็นหน้าที่ของกลไกทางแสงในตัวกล้อง ที่จะทำหน้าที่รับภาพเข้ามา



รูปที่ 2.5 ภาพที่สามารถมองเห็นได้จากกล้อง

2.4.3 การสุ่มเลือกจุดตำแหน่ง (ImageSampling)

เป็นการแปลงภาพ 2 มิติที่ได้ให้เป็นภาพเชิงดิจิทัล โดยการสุ่มเลือกทางจุดตำแหน่ง หรือ SpatiallySampling โดยสุ่มเลือกเฉพาะบางตำแหน่งในภาพ ซึ่งถ้าเราสุ่มเลือกมาละเอียดภาพที่ได้ก็จะมีรายละเอียดสูง หน่วยของการสุ่มเลือกก็คือ จุด หรือ Pixel ดังรูปที่ 2.6



รูปที่ 2.6 การสุ่มเลือกจุดตำแหน่งในภาพ

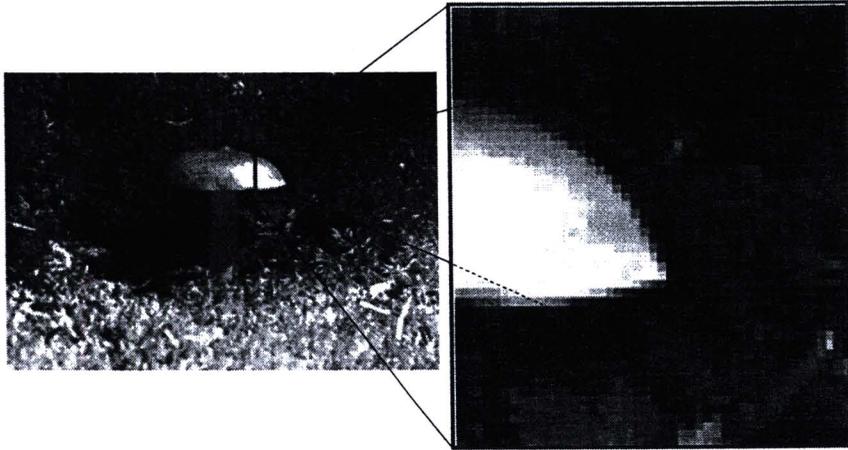
เหตุผลของการทำ ImageSampling เพราะว่า ในการแสดงผลของจอภาพคอมพิวเตอร์นั้น หน่วยของความละเอียดในการแสดงผลนั้นเป็นจุด (Pixel) แต่ในโลกแห่งความเป็นจริง ภาพที่มองเห็นด้วยตา ไม่ใช่การเรียงกันของจุด แต่มันเป็นภาพเชิงต่อเนื่องคือ ไม่สามารถแยกลงไปเป็นทีละจุด ๆ ได้เลย เพราะฉะนั้น เมื่อภาพนั้นมาอยู่ในคอมพิวเตอร์ ภาพจะต้องถูกปรับให้อยู่ในหน้าจอที่ประกอบด้วยจุดสี เพราะฉะนั้น จึงจำเป็นที่จะต้องมีการสุ่มเลือกจุดตำแหน่ง หรือการทำ ImageSampling จากที่ได้กล่าวมานี้ ถ้าเราสุ่มเลือกจุดตำแหน่งถี่มากเท่าใด คุณภาพของภาพที่ได้ก็จะดีขึ้น ลองดูเมื่อขยายเท่า ๆ กัน ดังรูปต่อไปนี้

ในการสุ่มเลือกทางตำแหน่งนี้ ถ้าในระยะความกว้างและความสูงของภาพ สุ่มมาละเอียดมาก ๆ เช่น สุ่มมา 128 จุด คุณภาพของภาพก็จะดีขึ้น แต่ก็ต้องใช้เนื้อที่ในการเก็บข้อมูลภาพมากขึ้น และขนาดของภาพก็จะมากขึ้นด้วย

2.4.4 การประมาณค่าความเข้มของแสง (ImageQuantization)

เมื่อได้ภาพจากการ Sampling มาแล้วแต่ละจุดในภาพจะถูกแทนด้วยสี ภาพในโทนสีเทา หรือ Grayscaleจะประกอบไปด้วยสีดำ และไล่เฉดสีจางลงไปจนถึงสีขาว
ดังรูปที่ 2.7





รูปที่ 2.7 การประมาณค่าความเข้มแสง

สีดำ จะแทนด้วยค่าตัวเลข 255 สีขาวจะแทนด้วยค่าตัวเลขคือ 0 รวมทั้งสิ้น 256 ระดับสี (0-255) หรือ 2 กำลัง 8 โดยที่ 8 ก็คือ จำนวนบิตในหน่วยความจำที่ใช้ในการเก็บค่านี้หนึ่งค่า เพราะฉะนั้น สีดำ จะถูกแทนด้วยรหัสในเลขฐานสองคือ 00000000 และสีขาวก็ จะถูกแทนด้วยรหัส 11111111 และสีที่อยู่ตรงกลางระหว่างสีดำกับสีขาวก็จะไล่ไปตามลำดับการนับของบิตในเลขฐานสองดังรูปที่ 2.8

	0	0000	0000
	1	0000	0001
	2	0000	0010
	3	0000	0011
	4	0000	0100
	-	-	-
	-	-	-
	254	1111	1110
	255	1111	1111

รูปที่ 2.8 การแทนค่าด้วยเลขฐานสอง

ถ้าภาพเป็นแบบโทนขาวดำ (Grayscale) แต่ละจุดภาพก็จะถูกแทนที่ด้วยตัวเลขที่บอกถึงค่าสีตั้งแต่ 0-255 ลองดูภาพต่อไปนี้ แต่ละจุดๆ จะถูกแทนที่ด้วยตัวเลข ซึ่งตัวเลขเหล่านี้ก็อยู่ระหว่าง 0-255 คือตั้งแต่ 0,1,2,3,4,..... 255 เป็นโทนสีเทา แต่ถ้าเป็นภาพขาวดำล่ะก็จะมีอยู่ด้วยกันแค่ 2 สีคือ สีดำ แทนด้วยเลข 0 กับสีขาวแทนด้วยเลข 255 เพราะฉะนั้นถ้าเป็นภาพขาวดำหนึ่งจุดภาพจะใช้พื้นที่เก็บข้อมูลเพียง 1 บิตเท่านั้น แต่ถ้าเป็นภาพในโทนสีเทานั้น ใน 1 จุดภาพจะใช้พื้นที่ในการเก็บข้อมูล 8 บิต ที่เป็น 8 บิตก็เพราะว่าค่าระดับสีเมื่อเปลี่ยนเป็นเลขฐานสองแล้วจะได้ 8 บิตค่าระดับสี 0 ก็คือ 00000000 ค่าระดับสี 1 ก็คือ 00000001 ค่าระดับสี 2 ก็คือ 00000010 ...ค่าระดับสี 256 ก็คือ 11111111 เพราะฉะนั้นใน 1 จุดสี จะมีค่าใดค่าหนึ่งเท่านั้น จึงใช้พื้นที่ในการเก็บเพียง 8 บิต บางครั้งจะใช้การเขียนในแบบยกกำลังคือ ภาพขาวดำ 2 ยกกำลัง 1 (1 ก็คือ 1 บิต) ภาพโทนสีเทา 2 ยกกำลัง 8 (8 ก็คือ 8 บิต)

ระดับความเข้มของสีที่เรากำหนดให้กับภาพ ถ้าภาพ ๆ นั้นใช้ระดับสีน้อย เช่น 256 , 128 , 64 , 32 , 16 , 8 , 4 , 2 (1 บิต) ผลที่ได้ก็คือการไล่สีของภาพนั้นจะไม่แนวล เป็นรอยที่ไม่ค่อยน่าดูเท่าไร ดังรูปที่ 2.9



ภาพต้นฉบับ



$$2^8 = 256 \text{ สี}$$



$$2^4 = 16 \text{ สี}$$



$$2^1 = 2 \text{ สี}$$

รูปที่ 2.9 ระดับความเข้มของสี

ในความเป็นจริงแล้ว ความละเอียดของภาพจะมีมากกว่านี้ และอัตราในการสุ่มภาพนั้น ถ้าไม่คำนึงถึงคุณภาพและรายละเอียดของภาพ แต่คำนึงถึงเนื้อหาของภาพ ใช้

อัตราน้อยๆ ได้ เช่น ต้องการทราบว่ารูปนี้คือรูปอะไรเท่านั้น เพราะฉะนั้น เมื่อพูดถึงเนื้อหาของภาพแบบดิจิทัล จะเกี่ยวข้องกับเรื่องของการมองเห็นของมนุษย์ด้วย เพราะภาพที่มีความละเอียดต่างกัน ดวงตาของมนุษย์อาจจะแยกแยะไม่ได้ หรือสนใจเพียงแค่นี้หา แต่ไม่สนใจในรายละเอียด ดังนั้นถ้ารูปๆ หนึ่ง sampling มาแค่ 2 ยกกำลัง 1 จะสามารถดูออกหรือไม่ว่าเป็นรูปหน้าคน และเพียงพอหรือไม่ที่จะนำไปใช้งาน ข้อมูลเพียงแค่ 2 บิตคือ 0 กับ 1 ถ้าเพียงพอและครอบคลุม แต่ถ้าต้องการข้อมูลมากกว่านี้ เช่น ทำระบบ Detect หน้าคนที่ต้องใช้ข้อมูลอื่น ๆ ด้วย เมื่อนั้นละ จึงต้องการข้อมูลจุดที่มากกว่า

2.5 ไครเรกต์เอกซ์ (DirectX)

ในอดีตการพัฒนาโปรแกรมทางด้าน Multimedia โดยเฉพาะการพัฒนาโปรแกรมเกมส์ ผู้พัฒนาโปรแกรมต้องเขียนโปรแกรมเพื่อติดต่อ Hardware โดยตรง ทั้งทางด้านภาพและเสียงทำให้ประสบปัญหาอย่างมาก เพราะต้องเขียนโปรแกรมที่รองรับ Hardware ที่หลากหลาย จึงได้มีการพัฒนาชุดคำสั่ง (application programming interfaces : API) สำหรับการติดต่อกับ Hardware เพื่อที่ผู้พัฒนาโปรแกรมเกมส์ไม่ต้องกังวลกับส่วนของการติดต่อกับ Hardware โดย API ที่เป็นที่ยอมรับคือ DirectX ของไมโครซอฟต์และ Open GL ของซิติคอนกราฟิกส์

ไมโครซอฟต์ได้พัฒนา DirectX พร้อมติดตั้งมากระบบปฏิบัติการตั้งแต่ Windows 95 เปิดให้ผู้ใช้งานทั่วไปสามารถ download เวอร์ชันใหม่มาปรับปรุงได้ฟรีจากเว็บไซต์ของ Microsoft ปัจจุบันพัฒนามาถึงเวอร์ชัน 10.0 ติดตั้งมาพร้อมกับระบบปฏิบัติการ Windows Vista

DirectX สามารถติดต่อกับระบบ Hardware ได้โดยตรง โดยไม่ต้องผ่านระบบปฏิบัติการ ทำให้สามารถเข้าถึงคุณสมบัติต่างๆ ของ Hardware เช่น Graphic Card , 3D Accelerator ได้โดยตรง และมีความสามารถวิเคราะห์ความสามารถของ Hardware เพื่อปรับแต่งค่าต่างๆ ของโปรแกรมให้เหมาะสมหรือชดเชยการทำงานที่ไม่สนับสนุนโดย Hardware นั้น โดยการจำลองการทำงานนั้นด้วย software ทำให้โปรแกรม Multimedia หรือเกมส์ที่ติดต่อผ่าน DirectX ใช้ความสามารถด้านฮาร์ดแวร์ของระบบได้อย่างเต็มประสิทธิภาพ และยังสามารถทำงานบนเครื่องคอมพิวเตอร์ที่มีระบบ Hardware ที่แตกต่างกันได้

2.5.1 ส่วนประกอบของ DirectX

2.5.1.1 DirectX Runtime

เป็นส่วนประกอบของ DirectX ทางฝั่งผู้ใช้โปรแกรมหรือผู้เล่นเกมส์ เพื่อให้โปรแกรม Multimedia หรือเกมส์สามารถทำงานกับ Hardware ได้ โดยปกติแล้วจะถูกติดตั้งมาพร้อมกับการติดตั้งระบบปฏิบัติการ Windows ถ้าเครื่องคอมพิวเตอร์เครื่องใดไม่มี DirectX Runtime ติดตั้งไว้ ก็จะไม่สามารถเล่นเกมส์ได้

หรือหากมีเวอร์ชันที่ไม่สอดคล้องกับที่โปรแกรมต้องการใช้ก็จะไม่สามารถทำงานได้ผู้ใช้ต้องติดตั้งDirectXให้ตรงตามที่โปรแกรมต้องการ

2.5.1.2 DirectX SDK

เป็นส่วนของผู้พัฒนาโปรแกรมโดย DirectX SDK เป็นเครื่องมือที่ใช้ในการพัฒนาโปรแกรมโดยไม่โครซอฟต์แวร์พัฒนา DirectX SDK สำหรับใช้การพัฒนาโปรแกรมในหลายภาษาเช่น C/C#, C# และ Visual Basic .NET โดยผู้พัฒนาโปรแกรมสามารถ Download ได้ฟรีจากเว็บไซต์ของ Microsoft โดยในชุดของ DirectX SDK จะมี DirectX Runtime ติดตั้งมาให้ด้วย

2.5.1.3 องค์ประกอบของ DirectX SDK

- Direct Graphics ทำหน้าที่จัดการเกี่ยวกับการแสดงผลทางจอภาพ ประกอบด้วยฟังก์ชันของ DirectDraw ที่สนับสนุนการสร้างภาพสองมิติ และDirect3D (D3D) สำหรับสร้างภาพสามมิติ
- Direct Audio ทำหน้าที่จัดการในการประมวลผลเสียงต่าง ๆ
- Direct Input ประกอบด้วยฟังก์ชันในการรับข้อมูลจากผู้ใช้ผ่านอุปกรณ์ต่าง ๆเช่นMouse, Keyboard และ Joystick เป็นต้น
- Direct Play ประกอบด้วยฟังก์ชันที่ใช้ในการเชื่อมต่อเครือข่าย (Network)
- DirectShow ใช้จัดการเกี่ยวกับมัลติมีเดียDirectSetupใช้เกี่ยวกับการ Install DirectX ใช้แสดงผลเวอร์ชันของ DirectX

2.6 ความรู้เกี่ยวกับไดเรกต์โชว์(DirectShow)

โดย DirectShow นับเป็นหนึ่งใน Component ของ DirectX ใช้ในการจัดการเกี่ยวกับมัลติมีเดีย เช่น ควบคุมการเล่น การหยุดเล่นของไฟล์มัลติมีเดีย เป็นต้น โดยที่ DirectShow นั้นทำงานอยู่บนพื้นฐานของ Filter graph ซึ่ง Filter graph นี้จะมีอยู่ 3 ชนิดด้วยกัน คือ

2.6.1 Source Filters

ใช้ทำหน้าที่ติดต่อกับไฟล์มัลติมีเดีย กล้องหรือwebcam เพื่อนำข้อมูลมาใช้

2.6.2 Transform Filter

ใช้ทำหน้าที่เปลี่ยนแปลงข้อมูลที่ได้รับมาจาก Source Filters ตามแต่ที่ผู้ใช้งานต้องการ เช่น เปลี่ยนข้อมูลมัลติมีเดียจากที่จะแสดงผลเป็นภาพสีให้เปลี่ยนเป็นขาวดำ

2.6.3 Render Filter

ใช้ทำหน้าที่แสดงผลข้อมูลมัลติมีเดียที่ได้รับมาโดยผู้ที่พัฒนาโปรแกรมที่จะทำงานร่วมกับกล้อง หรือwebcam จะต้องทำการเขียน Filter เหล่านี้ขึ้นมาตามแต่ผู้พัฒนาต้องการ

Filter แต่ละตัวจะมีส่วนที่เรียกว่า pin สำหรับเชื่อมต่อกับ filter ตัวอื่นกระบวนการทำงานของมันเกิดขึ้นจากการเลือกเอา Filter ทั้งสามแบบที่เหมาะสม มาต่อเข้าด้วยกัน อย่างเช่น เอา DV camcorder capture source Filter ต่อ DV Video Decoder transform filter และต่อเข้ากับVideo Renderer filter เพื่อจับภาพจากกล้อง DV ออกไปที่หน้าจอ

Directshow มีส่วนฉลาดนี้เพื่อเลือก filter ที่เหมาะสมมา connect เข้าด้วยกันให้เราอัตโนมัติโปรแกรมเมอร์จึงไม่ต้องกังวลว่า ควรจะเลือกเอา filter ตัวไหนมาต่อกันเราเพียงเอา source ใส่อเข้าไปมันก็จะเลือก filter และทำกระบวนการให้เองอัตโนมัติ

2.7 เกรย์สเกล(Grayscale)

ด้วยปกติภาพสีจะประกอบไปด้วยจุดหลายๆจุดรวมกันแล้วแต่ละจุดสีในรูปประกอบด้วยค่า R, G, B แต่ละค่ามีค่าระหว่าง 0-255 (256 ระดับ)

วิธีทำให้ภาพเป็นสีเทา คือ เอามาเฉลี่ยกัน $(R+G+B) / 3$ เช่น ตัวอย่างเป็นสีชมพู (255, 218, 178) ถ้าเราเฉลี่ยคือ ๆ จะได้เป็น $615 / 3$ ได้เป็น 217...เราก็ Plot สีลงไปโดยให้ค่า R, G, B เป็น 217 ทั้งสามค่า

วิธีทำก็จะวน Loop ทีละแถว ทีละจุดโดยวน $y = 0$ ถึง $\max Y$ แล้ววน $x = 0$ ถึง $\max X$ จากนั้นอ่านค่าสีที่จุดนั้นเสร็จแล้วจึงหาค่า $c = (R+G+B) / 3$ แล้วทำการ plot สีใหม่ลงไป จะได้ค่าที่เฉลี่ยได้อยู่ในช่วง 0-255 เหมือนเดิม เวลาเอาไปรวมให้เป็นสีโดย R, G, B เท่าๆ กัน ก็ไล่ Tone สีเทา แต่ถ้าอ่านค่าสีด้วย Point(x, y) แล้วซ้ำ ทางที่จะช่วยให้เร็วขึ้นคือใช้ API อ่านค่าสี, Plot จุด เช่นใช้ `GetBitmapBits()` โดยวนคำนวณค่าสีใหม่ทั้งรูปเหมือนกับการคำนวณ โดยทั่วไปจะมีตัวคูณเพื่อปรับค่า R, G, B ในภาพก่อน เช่น ตาม ITU standard

$$c = (222 * r + 707 * g + 71 * b) / 1000 \text{ หรือ } c = (0.11 * r) + (0.59 * g) + (0.3 * b)$$

2.8 ชุดคำสั่งภายใน (Aforge.NET)

เป็นฟังก์ชันหนึ่งของ Windows ที่สามารถเรียกใช้ได้จาก Visual C# โดยที่ไม่จำเป็นต้องเขียนโปรแกรมให้ยุ่งยาก เพียงแค่เรียกใช้ฟังก์ชันเท่านั้น ฟังก์ชันดังกล่าวจะประกอบไปด้วยฟังก์ชันย่อยอีกมากมาย แต่ฟังก์ชันที่เรียกใช้คือ `Aforge.Image` ซึ่ง `Aforge.Image` จะเป็นกรอบให้ตัวกรอง

แบบการประมวลผลภาพที่ใช้ในการกรองพิกเซลขึ้นอยู่กับค่าสีที่ต้องการ ตัวกรองการประมวลผลภาพเหล่านี้อาจนำไปใช้เพื่อให้พิกเซลซึ่งสีอยู่ภายในหรือภายนอกขอบเขตที่ระบุและกรอกข้อมูลส่วนที่เหลือของพิกเซลด้วยสีที่กำหนด

2.9 กรองสี (Color Filter)

การกรองสีการประมวลผลภาพของโปรแกรมจำเป็นที่จะต้องมีตัวกรองสีเพื่อกรองพิกเซลภายใน กับ ภายนอกโดยกำหนดช่วงสี RGB มันจะช่วยให้พิกเซลที่มีสีภายใน กับพิกเซลที่มีช่วงสีนอกช่วงที่กำหนดแยกออกจากกันอย่างชัดเจน

2.10 การสแกนับจุด (Blob Counter)

จะเป็นภาพที่ได้จากการประมวลผลทั้งหมดแล้ว จะเห็นเป็นภาพวัตถุกับพื้นหลัง โดยวัตถุอาจมีอยู่หลายจุด เพราะจะมีแสงรบกวนต่างๆเข้ามาเกี่ยวข้อง แต่ก็จะมีขนาดที่น้อยมากเนื่องจากเราทำการ Filter ทิ้งไปแล้ว จากนั้น Blob Counter จะทำหน้าที่นับวัตถุว่ามีกี่ชิ้น แล้วมีขนาดเท่าไร? สามารถเลือกได้ว่าเอาวัตถุที่มีขนาดเท่าไรขึ้นไปมาคิด ซึ่งในที่นี้สามารถกำหนดได้ว่า จะใช้กี่พิกเซล

2.11 เทรสโฮลด์ (Thresholding)

Thresholding คือค่าสำหรับแยกประเภทของพิกเซล การหาค่าเทรสโฮลด์นี้ได้ด้วยการพิจารณาฮิสโตแกรม (histogram) ของความเข้มของแสง(ฮิสโตแกรม คือกราฟที่แจกแจงจำนวนพิกเซลที่ความเข้มของแสงต่างๆ กัน)

2.12 การประมวลผลภาพเชิงตัวเลข (Digital Image Processing)

โดยปกติแล้วสายตาของคนทั่วไปจะมองเห็นภาพต่างๆเป็นลักษณะแบบอนาล็อก (Analog) ซึ่งสามารถอธิบายได้ด้วยคณิตศาสตร์ที่มีตัวแปรนับได้อย่างต่อเนื่องแต่เครื่องคอมพิวเตอร์จะใช้เลขฐานสองเป็นหลักในการคำนวณ โดยภาพจะถูกแทนที่ด้วยตัวเลขให้อยู่ในรูปของเมทริกซ์ (Matrix) ดังนั้นเมื่อนำภาพมาแปลงเข้าสู่เครื่องคอมพิวเตอร์ภาพนั้นจะกลายเป็นภาพดิจิทัล (Digital Image) ซึ่งจะเรียกการทำงานนี้ว่าเป็นการประมวลผลภาพเชิงตัวเลขระบบพิกัดระนาบ 2 มิติ (Spatial Coordinate) จะใช้ในการแสดงภาพดิจิทัลซึ่งมีขนาดความกว้างและความสูงของภาพแสดงในแกน X และแกน Y ส่วนจุดใดๆที่อยู่บนระนาบ XY จะเรียกแทนจุดใดๆนั้นว่าพิกเซล (Pixel) โดยจะมีฟังก์ชัน $f(x,y)$ ที่แสดงถึงค่าระดับความเข้มที่มีอยู่ L ระดับซึ่งเป็นจำนวนที่นับได้จำกัดแบบไม่ต่อเนื่อง (Discrete Quantity) ซึ่งค่าที่ได้เป็นผลมาจากการทำควอนไทต์ (Quantization) นั่นคือการ

แปลงจากอนาล็อกเป็นดิจิทัลจำนวนช่วงระดับความเข้ม L ของจุดภาพจะบ่งบอกถึงระดับความเข้มของภาพเชิงตัวเลขซึ่งโดยทั่วไประดับความเข้ม L จะมีค่าตั้งแต่ 2 ระดับขึ้นไปเช่นที่ค่าระดับความเข้ม 256 ระดับนั่นคือค่าระดับความเข้มของจุดภาพอยู่ในช่วง[0-255] โดยจะใช้เนื้อที่ในการเก็บขนาด 1 ไบต์ (Byte) หรือ 8 บิต(Bit) สำหรับข้อมูล 1 จุดภาพ ($28 = 256$) ในกรณีที่ต้องการภาพที่มีความละเอียดของระดับความเข้มสูงกว่านี้ขึ้นไปจะต้องใช้เนื้อที่ในการเก็บข้อมูลมากกว่า 8 บิต ซึ่งอาจจะเป็น 16 บิตหรือ 24 บิตโดยค่าระดับความเข้มของจุดภาพจะเป็น 216 ระดับและ 224 ระดับตามลำดับรูปแบบในการประมวลผลภาพสามารถแบ่งออกได้เป็น 2 ระดับคือการประมวลผลภาพในระดับต่ำ (Low-level Image Processing) และการประมวลผลภาพในระดับสูง (High-level Image Processing) ซึ่งมีรายละเอียดดังต่อไปนี้

การประมวลผลภาพในระดับต่ำเป็นการประมวลผลภาพเชิงตัวเลขเพื่อหาตัวแปรต่างๆมาอธิบายข้อมูลภาพโดยทั่วไปแล้วการประมวลผลภาพในระดับต่ำจะหมายถึงการประมวลผลภาพก่อน(Image Preprocessing) เช่นการกำจัดสัญญาณรบกวนการจำแนกภาพการหาขอบภาพการเปลี่ยนแปลงภาพการทำให้ภาพคมชัด เป็นต้นโดยการประมวลผลในระดับต่ำจะใช้ค่าความสว่างหรือระดับความเข้มของจุดภาพโดยตรงซึ่งถือว่าเป็นกระบวนการเบื้องต้นที่มีความสำคัญมากเพื่อนำตัวแปรที่ได้ไปใช้ในการประมวลผลภาพในระดับสูงต่อไป

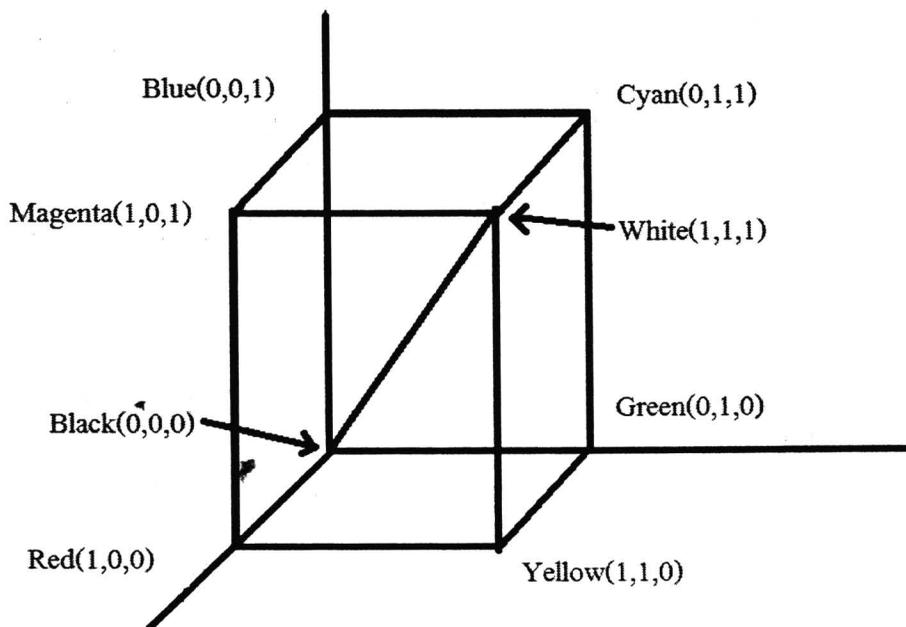
การประมวลผลภาพในระดับสูงเป็นการนำผลลัพธ์หรือตัวแปรที่ได้จากการประมวลระดับต่ำมาตีความหรือประมวลผลและแสดงให้อยู่ในรูปของสัญลักษณ์โดยสัญลักษณ์เหล่านี้จะแสดงถึงสิ่งต่างๆในภาพเพื่อให้คอมพิวเตอร์สามารถเข้าใจภาพได้เช่นการรู้จำรูปแบบของตัวอักษรความสัมพันธ์ระหว่างวัตถุในภาพขนาดหรือรูปร่างของวัตถุ เป็นต้น

2.13 แบบจำลองสี (Color Model)

สีเป็นคุณสมบัติทางกายภาพอย่างหนึ่งของวัตถุเกิดจากการรับรู้สัญญาณในรูปของคลื่นแสงที่ตกกระทบตามนุษย์และส่งสัญญาณผ่านประสาทตาไปยังสมองจากนั้นจึงแปลงค่าเป็นค่าสีต่างๆในระบบสีจะมีแบบจำลองสีหลายรูปแบบให้เลือกใช้ตามความเหมาะสมเช่นแบบจำลองสี XYZ ซึ่งเป็นแบบจำลองสีมาตรฐานที่พัฒนาขึ้นโดย Commission Internationale de l'Eclairage ถูกใช้เป็นตัวกลางในการอ้างอิงและแปลงค่าระหว่างแบบจำลองสีทั้งหมดในงานทางด้านกราฟิกจะใช้แบบจำลองสี CMY ในการแสดงภาพบนหน้าจอคอมพิวเตอร์จะใช้แบบจำลองสี RGB ในงานทางด้านคอมพิวเตอร์กราฟิกหรืองานที่ต้องคำนึงในด้านของศิลปะจะใช้แบบจำลองสี HSI หรือมาตรฐานในการวัดสีเช่นเครื่องมือในการวัดสีของผลิตภัณฑ์ทางการเกษตรจะใช้แบบจำลองสี $L^*a^*b^*$ เป็นต้นในงานวิจัยนี้ได้มีการศึกษาแบบจำลองสีต่างๆดังต่อไปนี้

2.13.1 แบบจำลองสี RGB

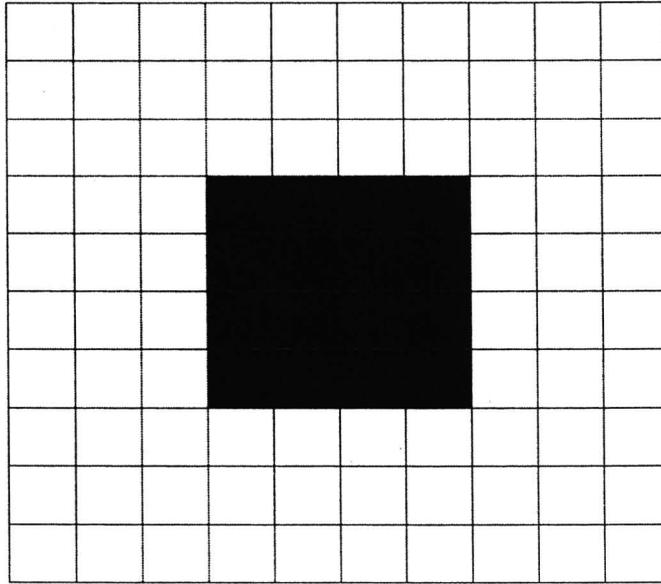
แบบจำลองสี RGB เป็นระบบสีที่ใช้กันทั่วไปจะประกอบไปด้วยแม่สีหลัก 3 สีคือ แดง (R)เขียว (G) และน้ำเงิน (B) ซึ่งแต่ละแม่สีจะมีขนาดเป็น 8 บิตดังนั้นในความคิดของแต่ละพิกเซลจะมีขนาดเป็น 24 บิตซึ่งในแต่ละแม่สีจะมีค่าตั้งแต่ 0 ถึง 255 ดังนั้นสามารถที่จะนำมาสร้างสีได้ถึง $256 \times 256 \times 256$ เท่ากับ 16,777,216 สีแบบจำลองสี RGB สามารถนำมาแสดงในระบบพิกัดปริภูมิสี (Color Space) ในแบบลูกบาศก์ได้ดังรูปที่ 2.10



รูปที่ 2.10แบบจำลองสี RGB

2.14 การวิเคราะห์ห้วงค์ประกอบที่เชื่อมต่อกันของภาพ (Connected Components)

ในภาพๆหนึ่งอาจจะมีองค์ประกอบหรือวัตถุอยู่วัตถุเดียวหรือมากกว่าหนึ่งวัตถุขึ้นไป โดยในแต่ละวัตถุอาจจะอยู่ในตำแหน่งที่ต่อเนื่องกันซ้อนทับกันหรือแยกออกจากกันซึ่งในภาพที่เป็นแบบไบนารีจะมีการพิจารณาการเชื่อมต่อของวัตถุโดยพิจารณาจากพิกเซลที่เป็นส่วนของวัตถุกับพิกเซลรอบข้าง (Neighborhoods) ว่ามีการเชื่อมต่อกันหรือไม่ซึ่งสามารถพิจารณาได้ทั้งใน 4 ทิศทาง(4-connected Neighborhoods) หรือใน 8 ทิศทาง (8-connected Neighborhoods) ซึ่งถ้าพิกเซลที่เป็นส่วนของวัตถุมีการเชื่อมต่อกันจะถือว่าเป็นวัตถุเดียวกันแต่ถ้าพิกเซลที่เป็นส่วนของวัตถุแยกออกจากกันหรือไม่เชื่อมต่อกันจะกำหนดให้เป็นวัตถุที่ต่างกันดังแสดงในรูปที่ 2.11



รูปที่ 2.11 องค์ประกอบของภาพ

จากรูปที่ 2.12 นั้นกำหนดให้พิกเซลสีขาวเป็นส่วนของพื้นหลังและพิกเซลสีดำเป็นส่วน
ของวัตถุเมื่อนำภาพมาวิเคราะห์องค์ประกอบที่เชื่อมต่อกันของภาพทำให้สามารถระบุหมายเลข
(Labeling of Connected Components) ให้กับองค์ประกอบของภาพได้ดังรูปที่ 2.12

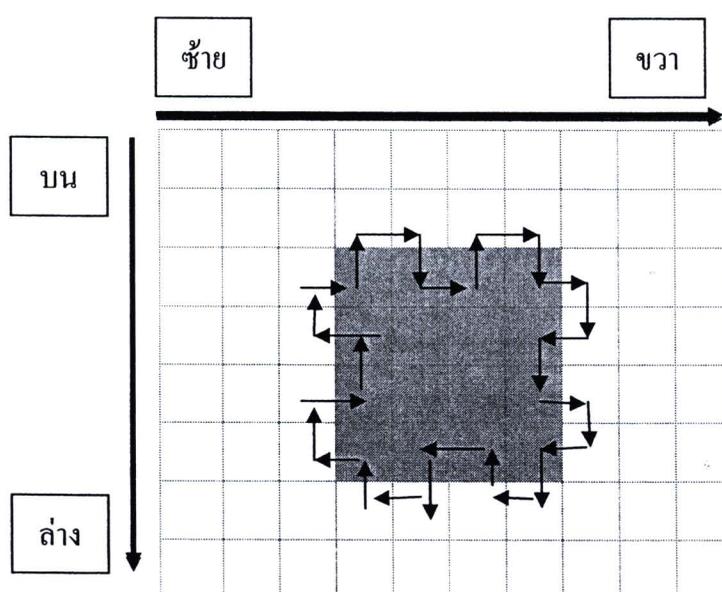
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	1	1	1	1	0	0	0
0	0	0	1	1	1	1	0	0	0
0	0	0	1	1	1	1	0	0	0
0	0	0	1	1	1	1	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

รูปที่ 2.12 การระบุหมายเลขให้กับองค์ประกอบของภาพโดยพิจารณาใน 4 ทิศทาง

จากรูปที่ 2.12 เป็นผลการระบุหมายเลขให้กับองค์ประกอบของภาพซึ่งภาพจะประกอบไปด้วย 4 องค์ประกอบด้วยกันโดยหมายเลข 0 แสดงส่วนของพื้นหลังหมายเลข 1 แสดงวัตถุที่หนึ่ง หมายเลข 2 แสดงวัตถุที่สองและหมายเลข 3 แสดงวัตถุที่สามของภาพซึ่งจากผลที่ได้ข้างต้นเป็นการวิเคราะห์องค์ประกอบที่เชื่อมต่อกันของภาพโดยใช้ 4 ทิศทางแต่ถ้าเป็นการวิเคราะห์โดยใช้ 8 ทิศทางจะทำให้ภาพข้างต้นประกอบไปด้วย 2 วัตถุเท่านั้นนั่นคือวัตถุที่หนึ่งและวัตถุที่สองเป็นวัตถุที่เชื่อมต่อกันในแนวทแยงจึงถือว่าเป็นวัตถุเดียวกัน

2.15 การติดตามเส้นขอบ (Contour Following)

การติดตามเส้นขอบเป็นการติดตามรอยขอบของวัตถุในภาพซึ่งการได้ไปตามขอบโดยใช้ความแตกต่างของสีจุดภาพระหว่างวัตถุกับพื้นหลังซึ่งโดยทั่วไปแล้ววิธีการนี้จะใช้กับภาพไบนารี แทนจุดภาพดำเป็นส่วนของวัตถุและแทนจุดภาพขาวเป็นส่วนพื้นหลังดังนั้นในการทำงานจะตรวจกวาดไปทุกๆจุดของภาพ โดยเริ่มจากจุดภาพแรกที่มุมซ้ายบนของภาพและตรวจกวาดไปในทิศทางจากซ้ายไปขวาและเลื่อนจากบนลงล่างเมื่อตรวจกวาดมาพบจุดภาพดำจุดแรกซึ่งเป็นส่วนของวัตถุจะเปลี่ยนลักษณะการเคลื่อนที่ไปยังจุดภาพถัดไปใหม่โดยมีเงื่อนไขการเคลื่อนที่ต่อไปนี้ กรณีที่จุดภาพปัจจุบันเป็นจุดภาพของวัตถุให้เลี้ยวซ้ายแล้วเคลื่อนตรงไปยังจุดภาพข้างหน้าถัดไป และกรณีที่จุดภาพปัจจุบันเป็นจุดภาพของพื้นหลังให้เลี้ยวขวาแล้วเคลื่อนตรงไปยังจุดภาพข้างหน้าถัดไปทำการเคลื่อนที่จุดภาพที่พิจารณาเช่นนี้ไปเรื่อยๆจนกว่าจุดปัจจุบันเป็นจุดเดียวกันกับจุดที่เริ่มต้นของวัตถุวิธีการหาขอบเขตของภาพสามารถแสดงได้ดังรูปที่ 2.13



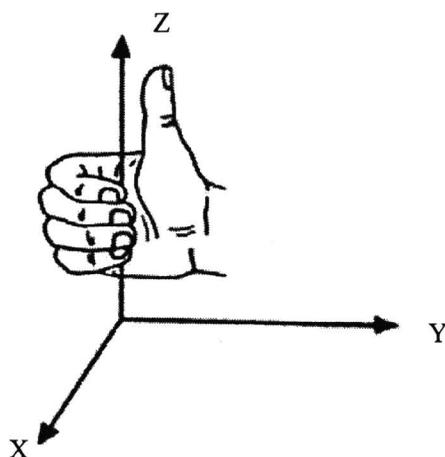
รูปที่ 2.13 วิธีการติดตามเส้นขอบ

2.16 ระบบพิกัดฉากในปริภูมิ 3 มิติ (Rectangular Coordinate in Three Dimension)

2.16.1 การกำหนดแกนพิกัด

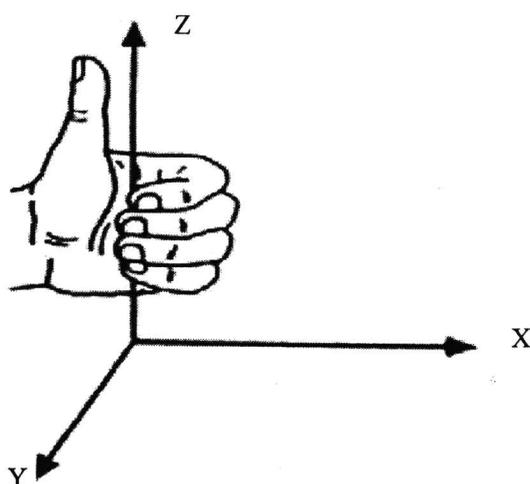
ในการกำหนดแกนพิกัดของระบบพิกัดฉาก 3 มิติ นั้นสามารถกำหนดโดยแบ่งออกได้เป็น 2 ระบบคือ

2.16.2 ระบบมือขวา (Right-handed System)



รูปที่ 2.14 การกำหนดแกนพิกัดฉาก 3 มิติ โดยระบบมือขวา

2.16.3 ระบบมือซ้าย (Left-handed System)



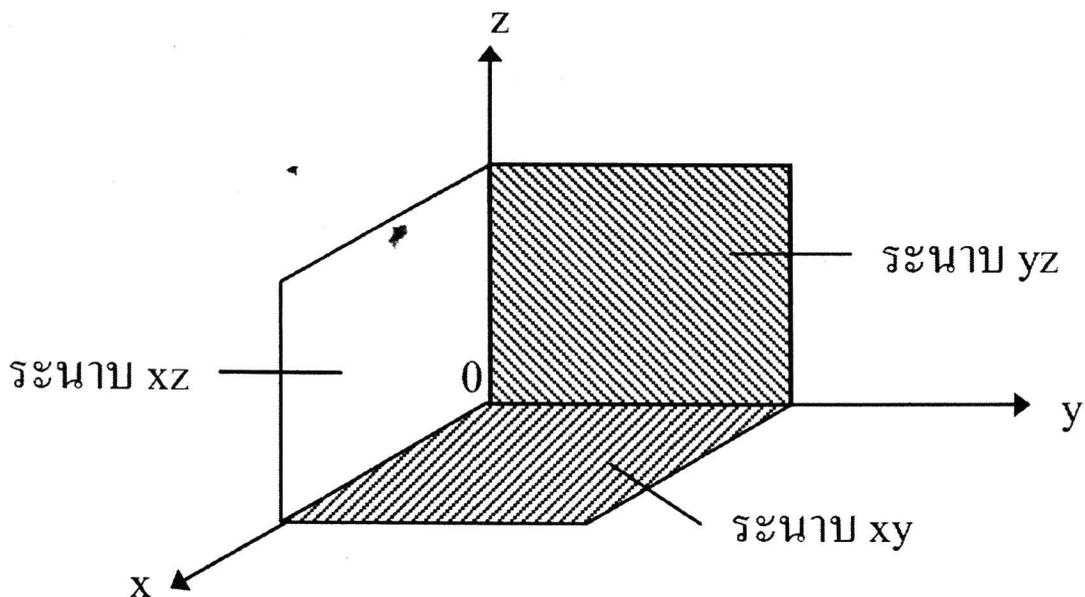
รูปที่ 2.15 การกำหนดแกนพิกัดฉาก 3 มิติ โดยระบบมือซ้าย

2.16.4 ระนาบพิกัดฉาก (Rectangular Coordinate Planes)

ในปริภูมิ 3 มิติ (Three Dimension) พิกัดของจุดหมายถึงระยะห่างจากระนาบ 3 ระนาบซึ่งต่างตั้งฉากซึ่งกันและกันและตัดกันที่จุดคงที่จุดหนึ่งซึ่งเรียกว่าจุดกำเนิด (Origin) ส่วนเส้นตรงทั้ง 3 เส้นที่ตั้งฉากกันเรียกว่าแกน X แกน Y และแกน Z โดยแต่ละแกนมีพิสัยบอกระยะห่างเท่าๆกัน แกน X แกน Y และแกน Z ที่ตัดกันทำให้เกิดระนาบ 3 ระนาบ เรียกว่าระนาบพิกัดฉากโดย

- ระนาบ XY มีแกน X และแกน Y อยู่ในระนาบ
- ระนาบ YZ มีแกน Y และแกน Z อยู่ในระนาบ
- ระนาบ XZ มีแกน X และแกน Z อยู่ในระนาบ

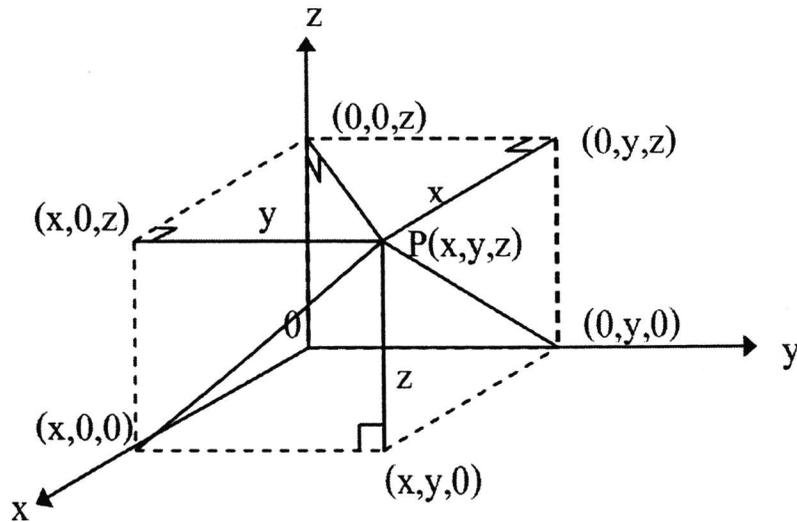
สามารถแสดงได้ดังรูปที่ 2.16



รูปที่ 2.16 ระนาบของปริภูมิ 3 มิติ

2.16.5 จุดพิกัด

ในการแสดงค่าพิกัดของจุดใดๆในปริภูมิ 3 มิติสามารถแสดงได้ดังรูปที่ 2.17



รูปที่ 2.17 ตำแหน่งพิกัดของจุดใดๆในปริภูมิ 3 มิติ

จากรูปที่ 2.17 ถ้าจุด P เป็นจุดใดๆในปริภูมิ 3 มิติเมื่อลากเส้นตรงจากจุด P ให้ตั้งฉากกับระนาบพิกัดฉากทั้งสามดังรูประยะห่างจากระนาบ YZ ถึงจุด P ซึ่งขนานกับแกน X เรียกว่าพิกัดแรก(X-coordinate) ของจุด P จะมีค่าเป็นบวกเมื่อจุด P อยู่ทางด้าน Ox ของระนาบ YZ และจะมีค่าลบเมื่อจุด P อยู่ฝั่งตรงข้ามเรียก Ox ว่าแกน X ทางบวกในทำนองเดียวกันพิกัดที่สอง(Y-coordinate)หมายถึงระยะห่างจากระนาบ XZ ถึงจุด P ซึ่งขนานกับแกน Y และพิกัดที่สาม (Z-coordinate) หมายถึงระยะห่างจากระนาบ XY ถึงจุด P ซึ่งขนานกับแกน Z ซึ่งพิกัดทั้งสามของจุด P จะเขียนแทนด้วย (x,y,z) เรียกว่าพิกัดฉาก (Rectangular Coordinate) ของจุด P

จุดทุกจุดในปริภูมิ 3 มิติจะมีพิกัดฉาก (x,y,z) เพียงพิกัดเดียวที่แทนจุดนั้นและในทางกลับกันทุกพิกัดฉาก (x,y,z) เมื่อ x, y และ z เป็นจำนวนจริงจะแทนจุดในปริภูมิ 3 มิติได้เพียงจุดเดียวเช่นกันด้วยวิธีการกำหนดพิกัดของจุดดังกล่าวทำให้สามารถจับคู่หนึ่งต่อหนึ่งระหว่างจุดในปริภูมิ 3 มิติกับพิกัด (x,y,z) เมื่อ x, y และ z เป็นจำนวนจริงได้

2.16.6 ประเภทกรอบพิกัด (Coordinate Frame)

ในการวิเคราะห์ความสัมพันธ์ระหว่างวัตถุในเชิง 3 มิติกับระนาบภาพในเชิง 2 มิตินั้นจะต้องมีกรอบพิกัดต่างๆที่นำมาใช้ในการพิจารณา

2.16.7 กรอบพิกัดวัตถุ (Object Coordinate Frame)

กรอบพิกัดวัตถุเป็นระบบพิกัดในเชิง 3 มิติจากรูปที่ 2.19จะระบุโดยใช้แกน x_o, y_o, z_o ซึ่งจะถูกนำมาใช้ในการแสดงแบบจำลองของวัตถุทั้งในงานทางด้านคอมพิวเตอร์

กราฟฟิกและคอมพิวเตอร์วิชัน โดยจะใช้ในการตรวจสอบหรือวิเคราะห์กับจุดต่างๆที่อยู่บนวัตถุ โดยตรงซึ่งพิกัดของวัตถุที่อ้างอิงด้วยกรอบพิกัดวัตถุจะไม่เปลี่ยนแปลงแม้ว่าวัตถุจะถูก หมุนหรือเคลื่อนย้ายตำแหน่งไปเช่นจุด B ที่วัตถุจะมีพิกัด $(x_b, 0, z_b)$ เป็นต้น

2.16.8 กรอบพิกัดโลก (World Coordinate Frame)

กรอบพิกัดโลกเป็นระบบพิกัดในเชิง 3 มิติจะระบุโดยใช้แกน x_w, y_w, z_w ซึ่งจะเป็นฉากที่ไว้สำหรับแสดงวัตถุหรือใช้ในการแสดงความสัมพันธ์ระหว่างวัตถุต่างๆหรือใช้ในการอ้างอิงการเปลี่ยนแปลงของวัตถุเช่นการหมุนการเปลี่ยนตำแหน่งเป็นต้นโดยอ้างอิงตำแหน่งเดิมและตำแหน่งที่เปลี่ยนไปจากระบบพิกัดโลก

2.16.9 กรอบพิกัดกล้อง (Camera Coordinate Frame)

กรอบพิกัดกล้องเป็นระบบพิกัดในเชิง 3 มิติ จะระบุโดยใช้แกน x_c, y_c, z_c ซึ่งจะถูกนำมาใช้ในการแสดงวัตถุที่สัมพันธ์กับตำแหน่งของกล้อง

2.16.10 กรอบพิกัดระนาบภาพ (Image Plane Coordinate Frame)

กรอบพิกัดระนาบภาพเป็นระบบพิกัดในเชิง 2 มิติ จะระบุโดยใช้แกน x_f, y_f ซึ่งจะถูกนำมาใช้ในการแสดงพิกัดของวัตถุที่ถูกฉายมายังระนาบภาพเช่นจากรูปจุด B ของวัตถุ จะถูกฉายมาบนระนาบภาพที่จุด b เป็นต้น