



สเตอริโอวิชั่น

STEREO VISION



ปฏิญานี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

ภาควิชาเทคโนโลยีการวัดคุมทางอุตสาหกรรม

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2534

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกฉบับที่ใช้

008407

ภาควิชา เทคโนโลยีการวัดคุมอุตสาหกรรม
สาขา เทคโนโลยีคอมพิวเตอร์อุตสาหกรรม
คณะ วิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้า เจ้าคุณทหาร ลาดกระบัง

เรื่อง สเตอริโอวิชั่น
(STEREO VISION)

ผู้จัดทำ

นาย พารา ลิ้มมะณีประเสริฐ 33.162122

นาย สมชาย แก้วกำแพง 33.162130

นาย อรุณชัย สัจจิตต์สวัสดิ์ 33.162137

.....อาจารย์ที่ปรึกษา

(อ. เกษตร์ ศิริสันติสัมฤทธิ์)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Thesis Title. STEREO VISION
Name : Mr. Para Limmaneeprasert 33.162122
Mr. Somchai Keawkumpang 33.162130
Mr. Anusit Songjitsawat 33.162137
Thesis Advisor. Kaset Sirisantsumrid
Level of study. Bachelor's Degree of Computer
Technology Instrument Department.
Academic Year. 1991

Abstract

This thesis describes stereo vision which uses information from two eyes to determine distance. Stereo vision works because it is easy to find the distance to something once the thing is found in two image from eyes with know separate. Stereo vision can be characterized by three step : (1) The point in one image corresponding to the projection of a point on a surface is located ; (2) The point in the other image corresponding to the projection of the same surface point is located ; and (3) The difference in projection of the corresponding point is used , together with estimates of the parameters of imaging geometry to determine a measure of the distance to the surface point.

สารบัญ

บทนำ

บทที่ 1 การมองเห็นของมนุษย์และคอมพิวเตอร์	1 - 1
(Human and Computer Vision)	
ระบบดิจิทัลอิมเมจโปรเซสซิง	1 - 8
(Digital Image Processing)	
-อิมเมจโปรเซสซิง	1 - 9
(Image Processing)	
-ดิจิทัลไลเซอร์	1 - 9
(Digitizer)	
-ดิจิทัลคอมพิวเตอร์	1 - 10
(Digital Computer)	
-สตอเรจดีไวซ์	1 - 10
(Storage Device)	
-ดิสเพลย์ดีไวซ์	1 - 11
(Display Device)	
บทที่ 2 รูปแบบของภาพและการโฟกัส	2 - 1
(Image Formation and Focusing)	
อัตราขยาย	2 - 3
(Magnification)	
ความยาวโฟกัส	2 - 5
(Focal length)	
ความลึก	2 - 8
(Depth of Field)	
ชนิดของเลนส์เมาต์	2 - 10
(Lens mounting)	
โซลิดสเตตคาเมอรา	2 - 11
(Solid State Camera)	
การประมวลผล	2 - 14
(Processing)	

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้สำหรับครูผู้ทำงานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3 การประมวลผลภาพเบื้องต้น	3 - 1
(Fundamental Concepts of Image Processing)	
พิกเซล	3 - 1
(pixel)	
วินโดว์	3 - 2
(Window)	
พิกเซลโลเคชัน	3 - 3
(pixel Location)	
เกรย์สเกล	3 - 6
(Gray Scale)	
ความผิดพลาดจากการตัดเศษ	3 - 10
(Quantizing Error)	
ความผิดพลาดจากการวัด	3 - 12
(Measurement Error)	
ฮิสโตแกรม	3 - 14
(Histogram)	
โครงสร้างของฮิสโตแกรม	3 - 16
(Construction)	
บทที่ 4 การประมวลผลการมองเห็นของเครื่องจักรเบื้องต้น	4 - 1
(Basic Machine Vision Processing)	
โมนาดิกวันพอยท์ทรานสฟอร์มเมชัน	4 - 3
(Monadic One Point Transformation)	
ไดอาดิกทูพอยท์ทรานสฟอร์มเมชัน	4 - 8
(Dyadic Two Point Transformation)	
คอนโวลูชัน	4 - 18
(Convolution)	
บทที่ 5 ขอบภาพ	5 - 1
(Edge Enhancement)	
ดิจิตอลฟิลเตอร์	5 - 1
(Digital Filter)	

โลว์พาสฟิลเตอร์	5 - 1
(Low Pass Filter)	
ไฮพาสฟิลเตอร์	5 - 4
(High Pass Filter)	
การหาขอบภาพ	5 - 8
(Edge Enhancement)	
บทที่ 6 การหาซีโร-ครอสซิง	6 - 1
(Zero-Crossing)	
ซีโร-ครอสซิงของลาปลาซเชียน-เกาส์เชียนฟิลเตอร์	6 - 1
(Zero-Crossing of Laplacian-Gaussian Filter	
Image)	
การตรวจจับการเปลี่ยนแปลงความเข้ม	6 - 3
(Detection of Intensity Change)	
ตำแหน่งของขอบภาพที่ถูกต้องที่ ซีโร-ครอสซิง	6 - 4
(Edge Position Accuracy at Zero-Crossing)	
การแยกตัวกระทำออกเป็นสองส่วน	6 - 10
(Separable Decomposition of LoG Operator)	
อัลกอริทึมพื้นฐานการทำนายเพื่อกำหนดตำแหน่งขอบภาพและทิศ	
ทางของขอบภาพ	6 - 11
(Predicate-Based Algorithm for Determining Edge	
Location and Orientation)	
บทที่ 7 การหาระยะทาง	7 - 1
(Finding Distance)	
การหาระยะทาง	7 - 1
(Range Finding)	
บทที่ 8 สเตอริโอวิชั่นและการทดลอง	8 - 1
(Stereo Vision and Experiment)	
สเตอริโอวิชั่นและการทดลอง	8 - 1
ผลการทดลอง	8 - 13
สรุปผลการทดลอง	8 - 16

กิตติกรรมประกาศ

หนังสืออ้างอิง

ภาคผนวก



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทนำ

เนื้อหาในปฏิญญาฉบับนี้ ในบทที่ 1 จะกล่าวถึงหลักการมองเห็นของมนุษย์และคอมพิวเตอร์ว่าเหมือนกันหรือแตกต่างกันอย่างไร ในบทที่ 2 จะกล่าวถึงอุปกรณ์ที่ใช้รับภาพและระบบเลนส์ของกล้อง ในบทที่ 3 จะเกี่ยวกับ การประมวลผลภาพในรูปแบบต่างๆ ในบทที่ 5-6 จะพูดถึงการหาขอบภาพของวัตถุ โดยวิธีการต่างๆ บทที่ 6 จะแนะนำวิธีการคำนวณหาระยะทางโดยวิธีของสเตอริโอวิชั่น ส่วนในบทที่ 7 จะเป็นการนำเอาทฤษฎีต่างๆ ตั้งแต่บทที่ 1-7 มาประยุกต์ใช้งานในการทำสเตอริโอวิชั่น



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1

การมองเห็นของมนุษย์และคอมพิวเตอร์

(Human and Computer Vision)

เทคนิคการใช้คอมพิวเตอร์ประมวลผลของรูปภาพ ได้ถูกนำมาพัฒนาและใช้ประโยชน์อย่างกว้างขวาง ถึงแม้ว่าเครื่องมือในการนำข้อมูลภาพ เข้ามาวิเคราะห์จะมีความยุ่งยากซับซ้อน โดยทั่วไประบบของ image-processing ในคอมพิวเตอร์มีขั้นตอนการทำงาน ซึ่งเน้นถึงการเกิดภาพ วิธีการนำภาพเข้าประมวลผล การประมวลผลเกี่ยวกับภาพและผลที่ได้ การบรรยายนี้เกี่ยวกับการเกิดภาพในลักษณะของฟิสิกส์ ซึ่งสัมพันธ์กับการมองเห็นของมนุษย์ ส่วนกระบวนการเกี่ยวกับภาพโดยคอมพิวเตอร์ จะเน้นเฉพาะส่วนการนำภาพมาหาขอบภาพ และนำมา match ภาพ เพื่อหาจุดเดียวกันของภาพทั้ง 2 เพื่อหาระยะทาง และความลึกของภาพ

จากการศึกษากระบวนการความสามารถแห่งการมองเห็นได้ของมนุษย์ ทำให้เกิดความเคลื่อนไหวในแนวทางที่จะพัฒนาการมองเห็นได้ของคอมพิวเตอร์ งานใดที่ทำได้ด้วยความสามารถของมนุษย์ย่อมเป็นการยากที่จะกล่าวว่างานนั้นไม่สามารถทำได้ด้วยเครื่องกล หรือ กล่าวอีกอย่างได้ว่าจากการศึกษาทางด้านกระบวนการของรูปภาพ ทำให้มีบทบาทสำคัญที่ให้ความหวังในการสร้างประสิทธิภาพของเครื่องคอมพิวเตอร์ ให้รับรู้สภาวะทางฟิสิกส์ได้มากขึ้น

ในเนื้อเรื่องส่วนนี้เป็นการกล่าวสั้นๆถึงคุณลักษณะของระบบการมองเห็น ผลของกระบวนการรูปภาพนี้ประเมินจากการสังเกตและตัดสินใจ ดังนั้นจึงเป็นสิ่งสำคัญที่สร้างระบบการมองเห็นเหมือนจริงบนรูปภาพที่ได้ ส่วนสำคัญเกี่ยวกับคุณลักษณะของการมองเห็นคือฟิสิกส์และสรีรศาสตร์

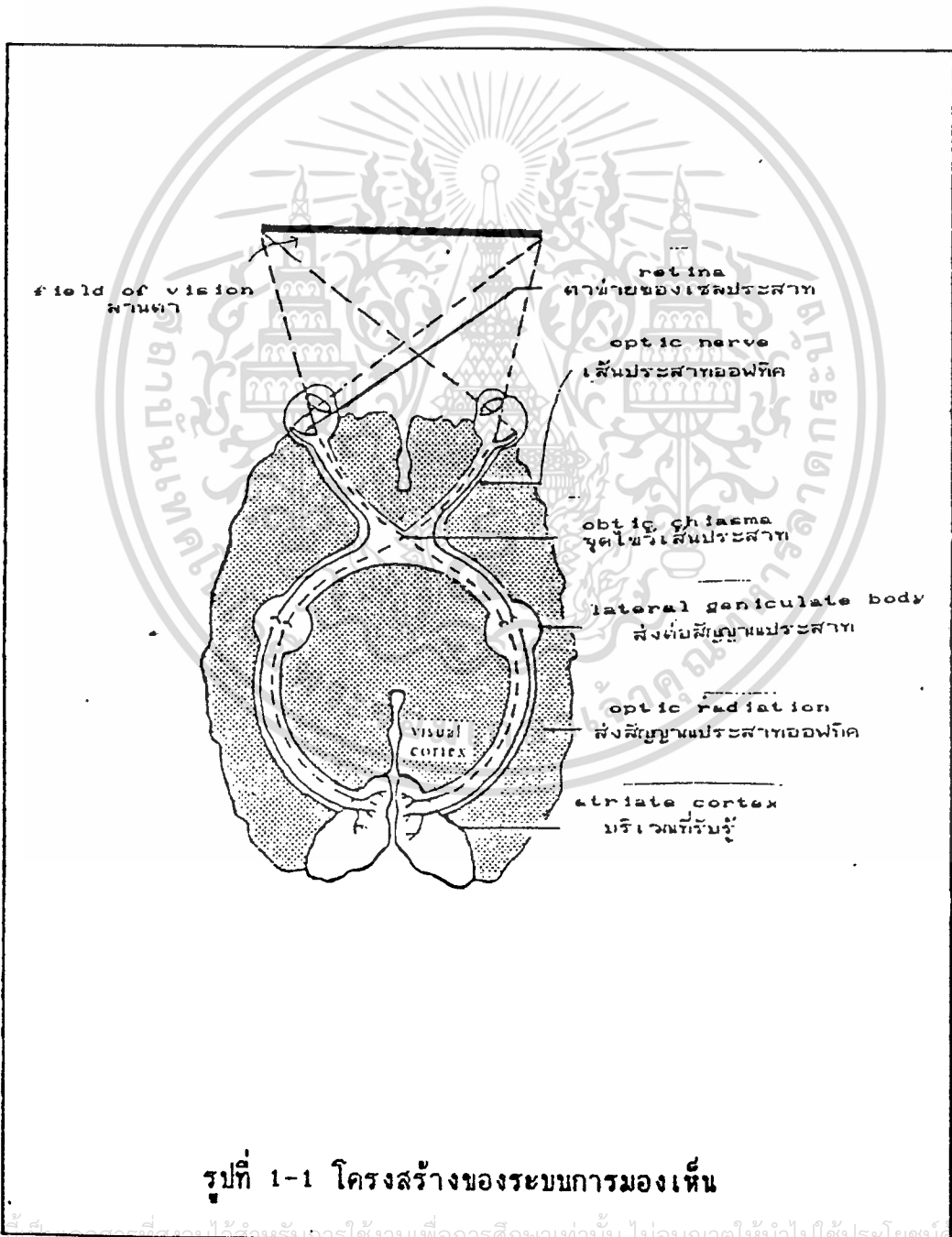
ระบบการรับภาพสามารถอธิบายได้ด้วยรายละเอียดทางกายวิภาคของระบบการมองเห็น แบ่งเป็นองค์ประกอบ 4 ส่วน คือ ดวงตา ประสาทเชื่อมโยงจากดวงตาส่วน lateral geniculate body และส่วนของ visual cortex ดังแสดงในรูป A จากปรากฏบนเรตินา ทั้งด้านซ้ายและขวา ผ่านกระบวนการทางเคมีภายในเปลี่ยนพลังงานแสงไปเป็นพัลส์ทางไฟฟ้า พัลส์เหล่านี้จะส่งผ่านไปสู่อptic chiasma ผ่าน lateral geniculate bodies และในที่สุดก็เข้าสู่บริเวณ visual cortex ใน opticital lobe ของสมอง

จากภาพตัดขวางดวงตามนุษย์ในรูป B เนื้อเยื่อที่อยู่นอกสุด sclera มีรูปร่างเกือบเป็นทรงกลมซึ่งมีรัศมีประมาณ 11 ม.ม. และความหนา 1 ม.ม. ที่ส่วนด้านหน้าของดวงตาเนื้อเยื่อ sclera จะกลมกลืนกันไปกับ cornea ซึ่งพองนูนมาด้านหน้าด้วยรัศมีประมาณ 8 ม.ม. หนา 1 ม.ม. ทางด้านหลัง optic nerve สอดเข้าไปในส่วน sclera จากทางด้านจุก ส่วน ciliary body อยู่บริเวณด้านหลังจุกรอยต่อระหว่าง cornea และ sclera ด้านหน้าของ ciliary body เป็น Iris ซึ่งเป็นช่องเกือบเป็นรูวงกลม และเป็นองค์ประกอบของแก้วตา

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

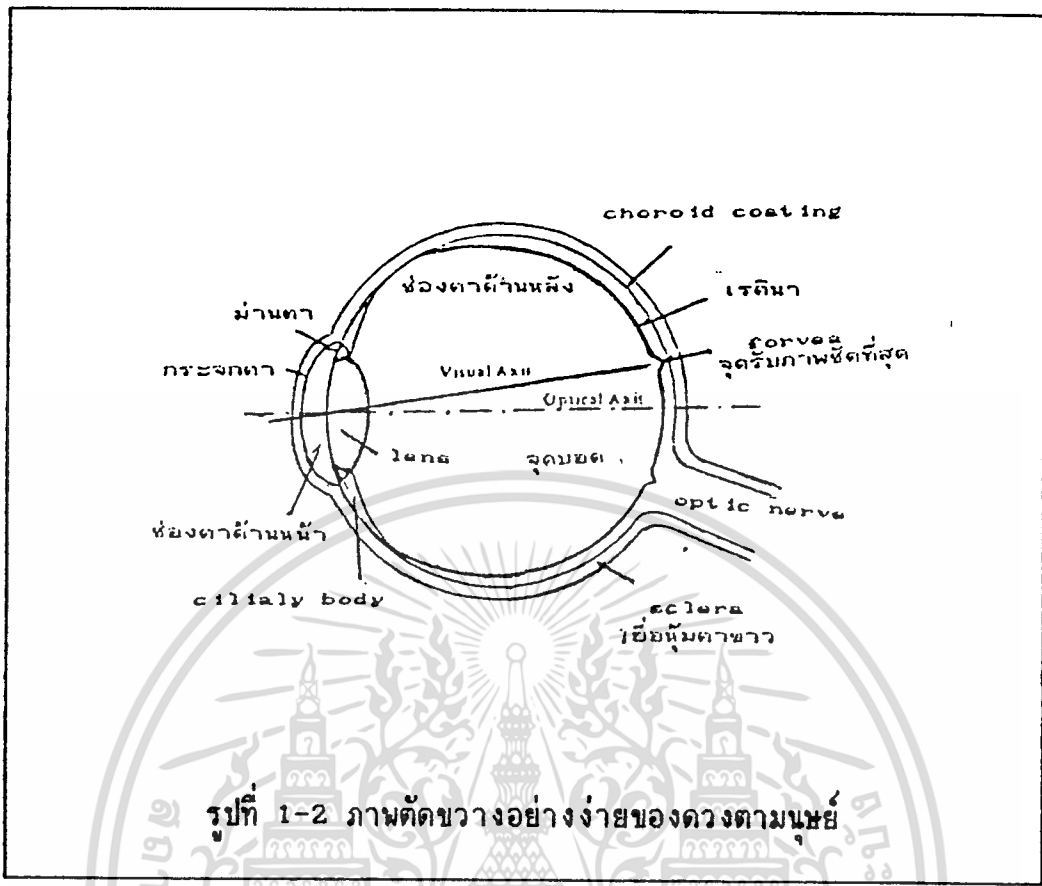
ขนาดการเปิดของแก้วตา เปลี่ยนแปลงได้ประมาณ 2 ถึง 8 ม.ม. โดยการหดและคลายตัวของกล้ามเนื้อที่ควบคุม Iris ปรากฏเป็นการตอบสนองความลึมน้ำระหว่างแสงและสรีรศาสตร์

ภายในแก้วตานั้นเลนส์นูนทั้งสองด้านลอยอยู่ในลักษณะที่เปลี่ยนรูปร่างได้ การเปลี่ยนรูปร่างได้นี้มีผลต่อการเปลี่ยนความยาวโฟกัสของเลนส์ ทำให้สามารถปรับระยะทางการมองเห็นได้ ช่องที่แยกจากกันโดยเลนส์จึงเกิดขึ้น anterior chamber เป็นส่วนที่มีของเหลวค้ำน้ำซึ่งอยู่ขณะที่ส่วนของ posterior chamber เต็มไปด้วยของเหลวที่มีความหนืดว ั้นในสุดของดวงตาปกคลุมด้วยส่วนที่มีความไวต่อแสงคือเรตินา

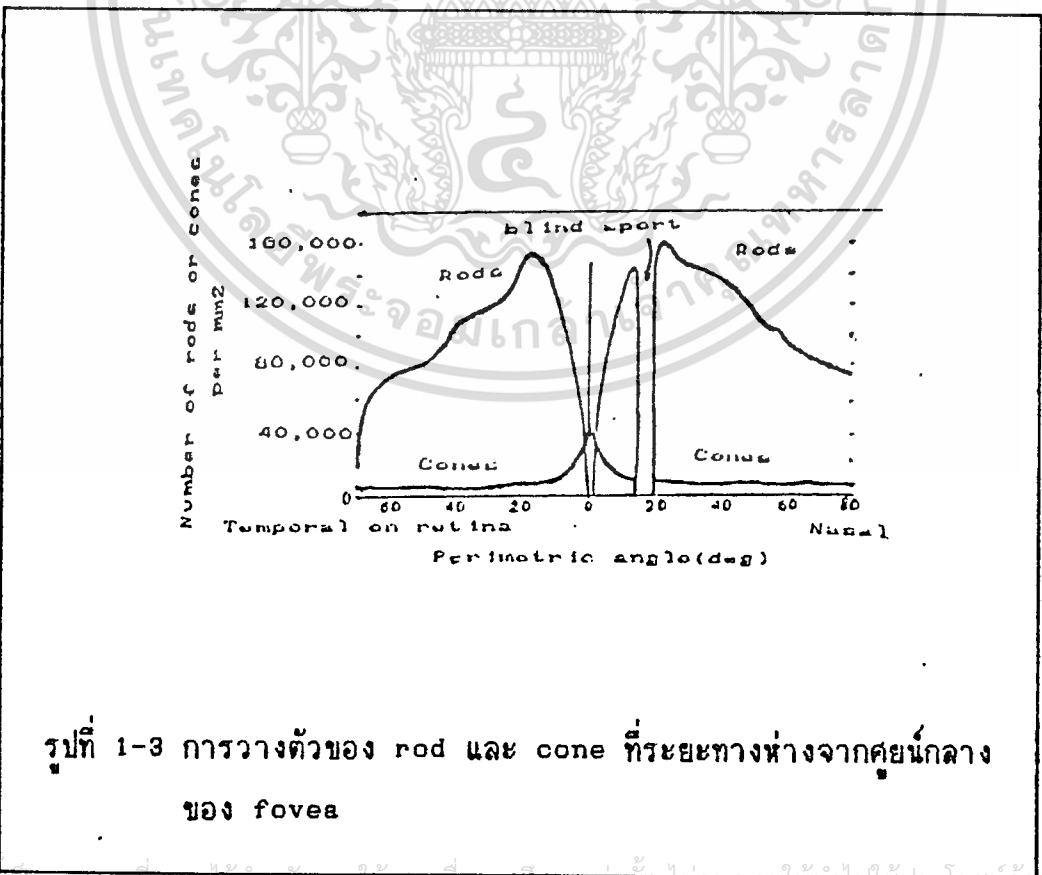


รูปที่ 1-1 โครงสร้างของระบบการมองเห็น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้ในวงที่ควรศึกษาเท่านั้น ไม่อนุญาตให้แก้ไขได้โดยไม่ขออนุญาต
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 1-2 ภาพตัดขวางอย่างง่ายของดวงตามนุษย์

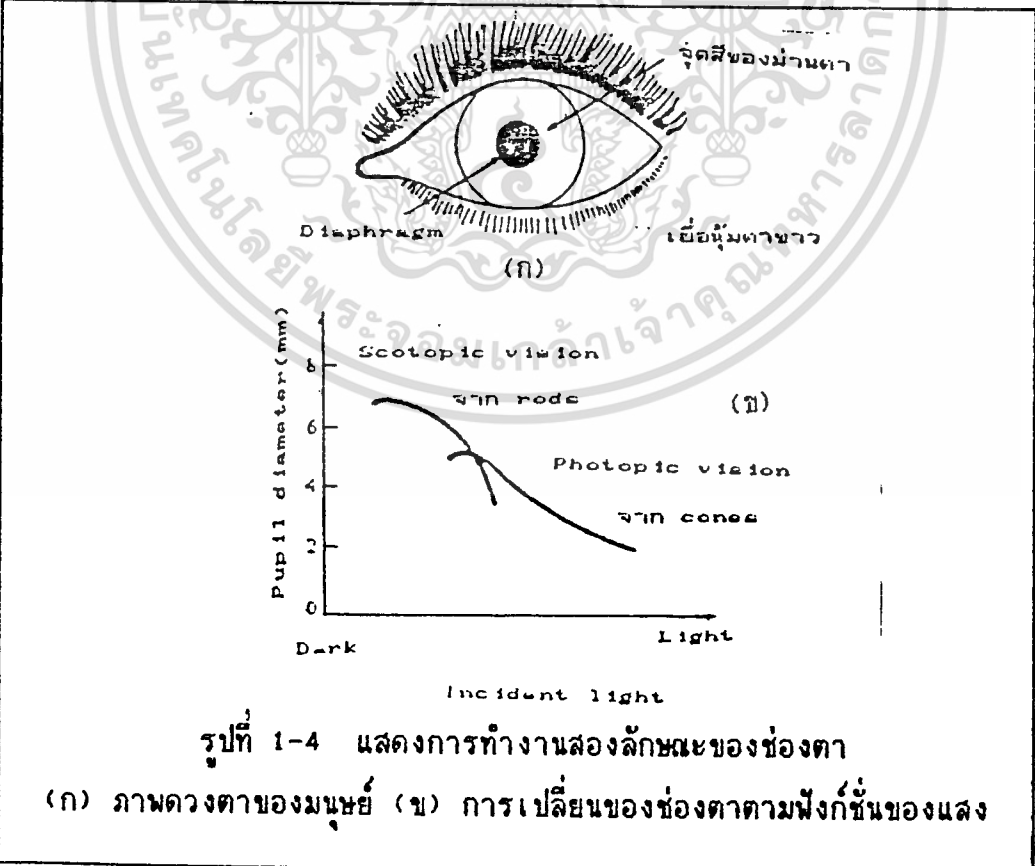


รูปที่ 1-3 การวางตัวของ rod และ cone ที่ระยะทางห่างจากศูนย์กลางของ fovea

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อการศึกษาเท่านั้น ไม่อนุญาตให้เผยแพร่ไปใช้ประโยชน์ในการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บนผิวเรตินาประกอบด้วยเซลล์รับแสงที่มีความหนาแน่นไม่เท่ากัน เรียกว่า "rod" และ "cone" การกระจายตัวของมันแสดงดังรูปที่ 1-3 ช่วงตรงกลางซึ่งมีความแคบประมาณ 0.4 มม. เรียกว่า "fovea" ประกอบด้วยส่วนที่เป็น cone เท่านั้น เนื่องจากตำแหน่งและมีความหนาแน่นของ cone ทำให้เป็นส่วนที่รับสีที่ดีที่สุด rod มีความหนาประมาณ 1-2 μm และยาว 50 μm สำหรับ cone มีขนาดแตกต่างกันไป แต่ส่วนที่อยู่บริเวณศูนย์กลางของ fovea มีความหนาประมาณ 1-2 μm เช่นเดียวกัน จำนวนของ cone มีประมาณ 5-7 ล้าน และจำนวน rod ประมาณ 75-150 ล้าน ในบริเวณ fovea การเชื่อมต่อระหว่าง cone กับ optical fibers ประมาณค่าเป็นหนึ่งต่อหนึ่ง เมื่อระยะทางจาก fovea เพิ่มขึ้น จำนวนของการเชื่อมต่อยิ่งมีค่ามากขึ้น rod จะมีความไวต่อความส่องสว่างต่ำและมีการตอบรับต่อ scotopic ส่วน cone ซึ่งมีความหนาแน่นมากในบริเวณ fovea อยู่ในตำแหน่งตามเส้นทางการมอง และตอบรับกับระบบการมอง photopic มากที่สุด

ช่องเปิดของลูกตาจะเปิดกว้างขึ้นเมื่อมีแสงสว่างปริมาณมากขึ้น เพื่อลดจำนวนแสงที่ตกกระทบสู่เรตินา อย่างไรก็ตามเมื่อเวลาผ่านไประยะหนึ่งเรตินาจะปรับตัวเข้าสู่ระดับใหม่ และช่องเปิดกลับเข้าสู่ระดับเดิม ช่องเปิดนี้สามารถควบคุมปริมาณแสงด้วยแฟกเตอร์ 30 เรตินานั้นสามารถปรับตามขนาดของแสงได้เช่นกันแต่ด้วยแฟกเตอร์ 10^{10}



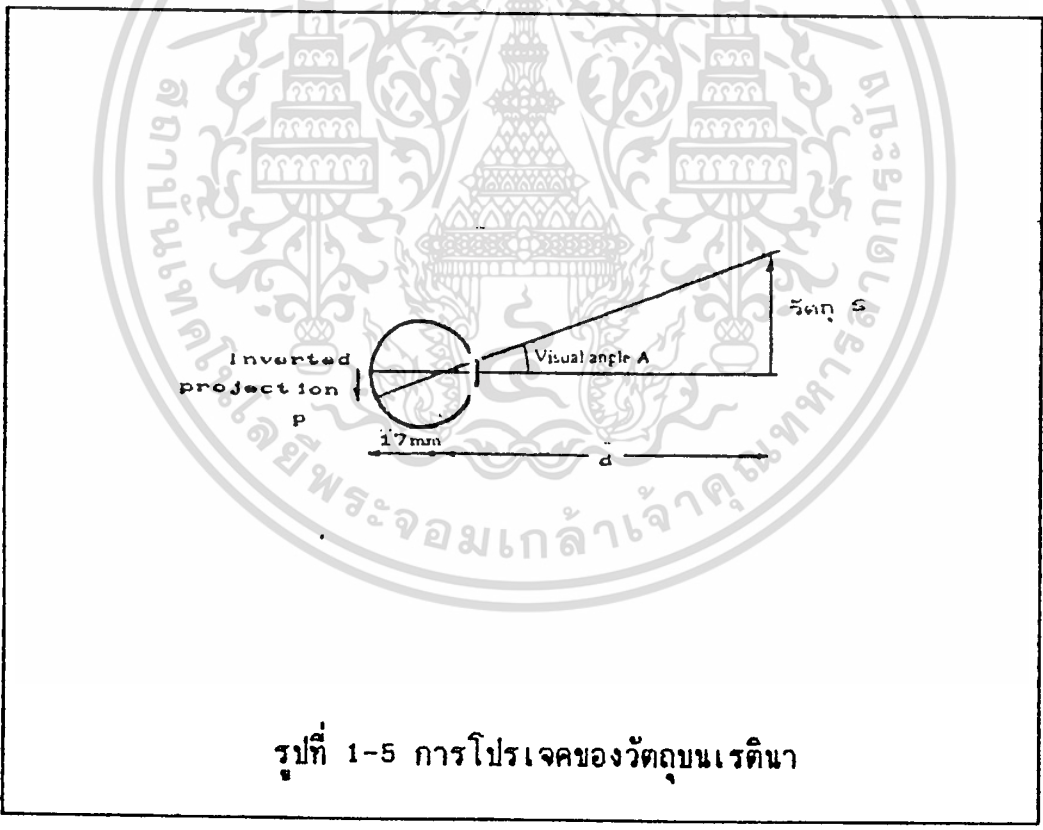
รูปที่ 1-4 แสดงการทำงานสองลักษณะของช่องตา
 (ก) ภาพดวงตาของมนุษย์ (ข) การเปลี่ยนของช่องตาตามฟังก์ชันของแสง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หลังจากแสงผ่าน aqueous humor มันจะผ่านช่องเปิดของตาบริเวณศูนย์กลางของม่านตา จากคุณสมบัติของเมคิลที่ม่านตาทำหน้าที่คล้ายเครื่องกลในการควบคุมการปิดเปิดช่องตา เป็นภาพด้านหน้าของดวงตามนุษย์ การเปลี่ยนแปลงขนาดของแผ่นปิดเปิดนี้ขึ้นอยู่กับกล้ามเนื้อสองชนิด คือ sphincter และ dilator pupillae กล้ามเนื้อ sphincter ใช้สำหรับการบีบรัดโดยเคลื่อนที่ขนานไปกับวงกลมของม่านตา และ dilator ใช้สำหรับการขยาย ขนาดของช่องว่างตามฟังก์ชันแสงตกกระทบนั้นแสดงดังรูป 1-4 (ข)

การเปลี่ยนแปลงขนาดของช่องตานั้นสนองต่อ 3 วัตถุประสงค์ คือผลต่อฟังก์ชันการสะท้อนแสง เพื่อปรับขนาดปริมาณของแสงที่เข้าสู่ดวงตา การปรับสายตาเพื่อรับวัตถุ และปรับภาพภายใต้เงื่อนไขจากแสงที่สว่างมาก

ภาพถูกปรับระยะการมองได้ โดยเปลี่ยนระยะโฟกัสให้ภาพตกลงที่เรตินา และผลความสูงของภาพกลับหัวบนพื้นเรตินาจะสัมพันธ์กับวัตถุดังรูป 1-5



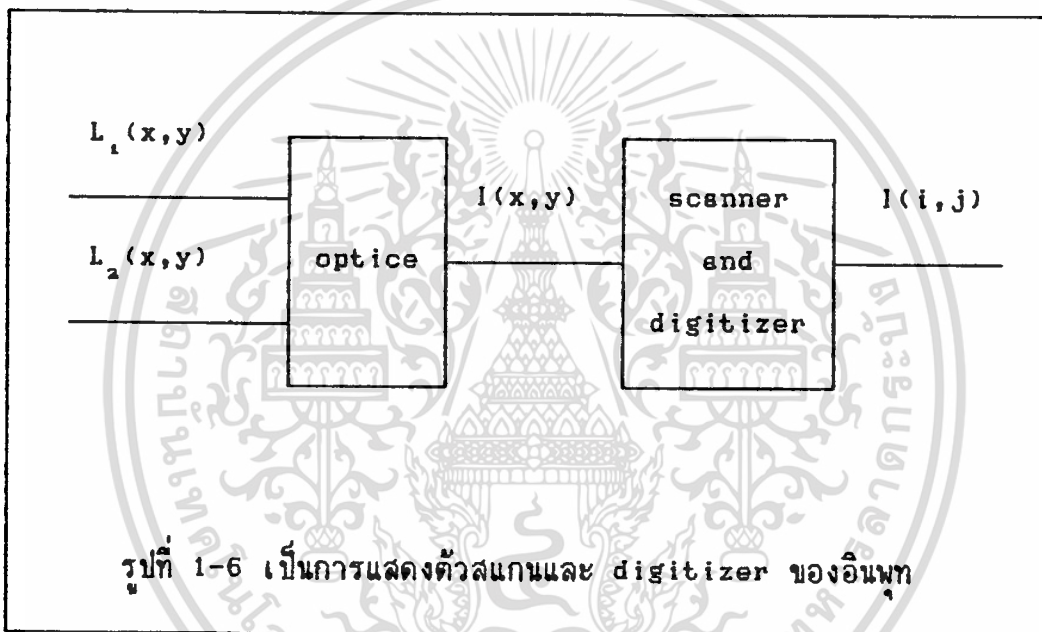
รูปที่ 1-5 การโปรเจคของวัตถุบนเรตินา

พารามิเตอร์ที่สัมพันธ์กับการมองเห็น และใช้ประโยชน์ในกระบวนการรูปภาพ เพื่ออธิบายให้เข้าใจในรายละเอียดคุณสมบัติเกี่ยวกับการมองเห็น จากแนวคิดพื้นฐานต่างๆไป พารามิเตอร์เหล่านี้คือ ความสว่าง ความชัด ความละเอียด และ ความคมชัด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Brightness เป็นแนวคิดทางจิตศาสตร์หรือเป็นความรู้สึกสัมพันธ์กับจำนวนของแสงที่ถูกสร้างขึ้น เนื่องจากดวงตามนุษย์มีความสามารถปรับตัวตามขนาดของแสงอย่างมาก ดังนั้น มนุษย์จึงไม่สามารถจะตัดสินใจเกี่ยวกับค่าปริมาณแสงสว่างได้อย่างถูกต้อง

Lightness นั้นสัมพันธ์กับการรับรู้ของผู้สังเกตของความแตกต่างใน ความเป็นสีค่าขาวหรือความเป็นสีเทาระหว่างวัตถุ ดังนั้นจึงแตกต่างจาก brightness คำว่า "contrast" ใช้สำหรับบ่งบอกลักษณะความแตกต่างในความสว่างนั้น ค่าความเป็นสีเทาขึ้นอยู่กับฉากเปรียบเทียบเบื้องหลัง



การแปลงภาพจริงเป็นข้อมูลภาพ

เราพิจารณาแหล่งพื้นฐานของข้อมูลภาพ 2 ลักษณะ กรณีแรกคือการสร้างสัญญาณของสิ่งมีชีวิตจากภาพจริง และในกรณีของการคืนภาพเดิมจากข้อมูลที่เก็บไว้ในตัวกลางทางฟิสิกส์ ถ้าเราสังเกตจากรูป 1-6 ที่ขั้นตอน optic ซึ่งจะสร้างสัญญาณภาพ $I(x,y)$ ป้อนให้กับส่วนของการสแกนและ digitizer เป็นส่วนของการนำข้อมูลของภาพเข้าสู่คอมพิวเตอร์ ภาพอนาล็อก $L_1(x,y)$ หรือ $L_2(x,y)$ ถูกเปลี่ยนเป็นภาพสัญญาณดิจิทัล $I(x,y)$ และ บางครั้งเรียกว่า เป็นส่วนเรตินาของคอมพิวเตอร์ เราจะอธิบายถึง hardware ในส่วนที่จำเป็นสำหรับการให้รูปภาพเป็นอินพุทของคอมพิวเตอร์เรากำหนดให้ภาพดิจิทัล $I(i,j)$ เป็นฟังก์ชันจริง 2 แกนของตัวแปร discrete ค่า i,j ซึ่งเป็นค่าอ้างอิงถึงแสงและเงา สีระดับที่ปรากฏบนภาพ ค่าเหล่านี้สัมพันธ์กับอินพุท $L_1(x,y)$ หรือ $L_2(x,y)$ อย่างไรอย่างหนึ่งในบริเวณ (x,y) ในบริเวณนี้เป็นค่า discrete ของจุดต่างๆบนภาพและเรียกว่า pixel ในเรื่องของ hardware มีการทำงานสองชั้นไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

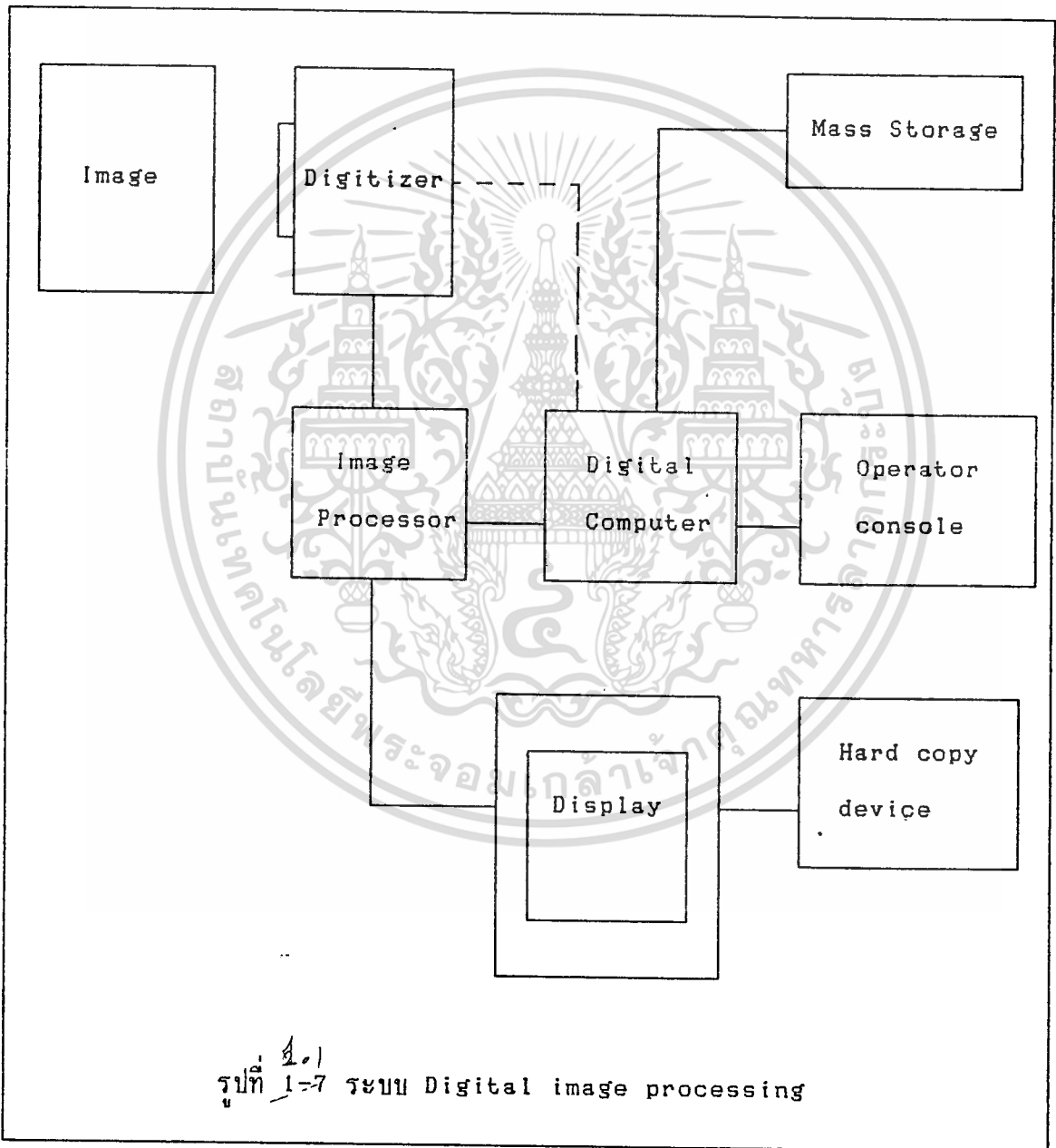
ตอน คือส่วนของการสุ่มข้อมูลจากภาพเป็น pixels และการทำ quantize ของระดับสีเทา ขนาดแนวกว้างยาวของภาพจะเป็น 128 x 128, 256 x 256 และ 512 x 512 แต่เราอาจขยายขึ้นถึง 2048 x 2048 ซึ่งมักจะไม่ใช่ในงานปกติของ remote sensing ระดับของ gray level นั้นมีระดับความแตกต่างกันเพียง 64 ระดับก็สามารถให้ความสมบูรณ์พอเพียง ถึงแม้ว่าในบางครั้งมีการใช้ถึง 256 ระดับ



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ระบบ Digital Image Processing

ส่วนประกอบพื้นฐานของระบบ digital image processing แสดงในรูป 1-7
การทำงานของแต่ละ block ในรูป 1-7 อธิบายได้ดังนี้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



๕.๒ Image Processors

digital image processor เป็นหัวใจของระบบ image processing ซึ่งระบบ image processor ประกอบด้วย ชุดของ Hardware ซึ่งทำหน้าที่พื้นฐาน 4 อย่างคือ

1. การรับภาพ จะมีสัญญาณ T.V. เป็น อินพุท และจะเปลี่ยนสัญญาณนี้ไปเป็นสัญญาณดิจิทัล ปัจจุบันนี้ อิมเมจ โพรเซสซิ่ง (image processing) โดยมาก สามารถรับภาพจาก T.V. ได้ใน 1 frame-time (ใช้เวลา 1/30 วินาที) เพราะว่าภาพ ที่รับเข้ามามักจะเป็น "frame grabber"
2. การเก็บภาพ เราเรียกว่า frame buffer เป็นหน่วยความจำที่สามารถนำภาพเข้ามาเก็บไว้ได้และยังสามารถที่จะโหลด (Load) หรือ อ่านจากสัญญาณ T.V. (เพื่อที่จะได้ภาพต่อวินาที)
3. การโปรเซสซิ่ง ในระดับล่าง เป็นการนำภาพที่ถ่ายเข้ามาทำการประมวลผลและการกระทำทางลอจิก (logic) ซึ่งตัวกระทำนี้เราเรียกว่า Arithmetic - Logic Unit (ALU) ซึ่ง ALU เป็นอุปกรณ์ทาง ฮาร์ดแวร์ ที่ถูกออกแบบมาเป็นพิเศษ ที่ใช้ในความเร็วสูงและเป็นการโปรเซสซิ่ง ในแบบขนาน
4. การแสดงผล จะทำการอ่านจากหน่วยความจำที่เก็บภาพไว้แล้ว และทำการเปลี่ยนข้อมูลดิจิทัลไปเป็นสัญญาณ อนาล็อก และ เอาท์พุท ของสัญญาณนี้จะส่งไปที่ T.V. มอนิเตอร์ หรือ อุปกรณ์ วิดีโอ อื่นๆ

Digitizers

Digitizer จะทำการเปลี่ยนภาพไปเป็นความเข้มของแสง (เป็นตัวเลข) ตามความเหมาะสมของสัญญาณอินพุทที่รับเข้ามา ซึ่งเป็นดิจิทัล โดยมากอุปกรณ์เหล่านี้จะเป็น micro-densitometers, flying spot scanners, image dissectors, vidicon cameras and photosensitive solid-state arrays. ส่วน Vidicon cameras สามารถรับภาพของธรรมชาติ เข้ามาได้ โดยขึ้นอยู่กับความเข้มของแสงที่ทำการรับเข้ามาได้

Digital Computers

อิมเมจ โปรเซสซึ่ง ถูกจัดมีความสามารถในการ โปรเซสซึ่ง ภายในดังนั้น สิ่งหนึ่งมักจะพบว่า อิมเมจ โปรเซสซึ่ง จะเชื่อมต่อกับทุกๆส่วนของ คอมพิวเตอร์ ซึ่งจัดให้แต่ละ โปรแกรมสามารถเข้าไปประมวลผลข้อมูลของ อิมเมจ ได้ ระบบคอมพิวเตอร์ ที่ใช้ในการ โปรเซสซึ่ง ของข้อมูลของ อิมเมจนั้น มีตั้งแต่เครื่อง ไมโครคอมพิวเตอร์ ไปจนถึง ระบบ คอมพิวเตอร์ ขนาดใหญ่ ซึ่งมีความสามารถในการคำนวณบนพื้นที่ของ อิมเมจ ขนาดใหญ่ ตัวแปรสำคัญที่มีอิทธิพลต่อโครงสร้างของ คอมพิวเตอร์ สำหรับใช้ในการ โปรเซสซึ่ง ของข้อมูล อิมเมจ นั้น จะเป็นจุดมุ่งหมายในการเอาไปประยุกต์ใช้มากกว่า และความต้องการของการใช้งานของข้อมูลการนำไปใช้กับเครื่อง ไมโครคอมพิวเตอร์ หรือ เครื่อง มิโนคอมพิวเตอร์ ก็เพียงพอแล้ว

Storage Devices

ดิจิตอล อิมเมจ (digital image) ซึ่งมีขนาด 512 x 512 พิกเซล ซึ่งแต่ละพิกเซล ต้องการพื้นที่ 8 บิต ดังนั้นภาพทั้งภาพจะใช้เนื้อที่ประมาณ 0.25 เมกะไบต์ การเก็บภาพส่วนใหญ่จะมีอุปกรณ์อยู่ 3 ชนิดคือ magnetic disk, magnetic tapes และ optical disk

Magnetic disks จะมีความจุประมาณ 700 เมกะไบต์ หรือมากกว่านั้น ดังนั้น 700 เมกะไบต์ ของ ดิสค์ ควรจะเก็บภาพได้ 2800 ภาพ (ภาพมีขนาด 512x512)

Magnetic tapes ที่มีความเข้มสูง (6400 ไบต์ต่อนิ้ว) สามารถเก็บภาพ 1 ภาพโดยใช้พื้นที่ใน เทป ประมาณ 4 ฟุต

Optical disks ซึ่งเป็นการนำเทคโนโลยีการอ่านและเขียนโดย เลเซอร์ ความจุในการเก็บของ single optical disk platter มีประมาณ 4 จิกะไบต์ หรือเก็บเป็นภาพได้ประมาณ 16,000 ภาพ

Display Devices

- มอนิเตอร์ แบบ โมโนโครม และ สี เป็นอุปกรณ์แสดงผลที่ใช้ในระบบ อิมเมจ โปรเซสซึ่ง ในปัจจุบัน มอนิเตอร์ จะถูกขับโดยเอาท์พุท ของ image display module ใน อิมเมจ โปรเซสซึ่ง ซึ่งจะส่งสัญญาณออกที่ เอาท์พุท เป็นสัญญาณภาพไปที่ จอภาพของมอนิเตอร์ ส่วนเครื่องแสดงผลอย่างอื่น ๆ จะรวมถึงจอภาพ CRT และ Printer ด้วย
- ระบบจอภาพแบบ CRT ตำแหน่งของ อิลิเมนต์ ต่างๆ ทางแนวนอนและแนวตั้งของภาพ เป็นการเปลี่ยนแปลง โวลเตจ ที่นำไปจ่ายให้กับ โยค คอยล์ (Yoke coil) เพื่อใช้ควบคุมการเบี่ยงเบน Electro Beam ของ CRT ทั้งแนวตั้งและแนวนอน
- ระบบ Printer เป็นอุปกรณ์แสดงผลเบื้องต้นสำหรับงาน อิมเมจ ที่มีความละเอียดต่ำๆ ระดับ gray level ของจุดต่างๆ ที่พิมพ์ออกมาสามารถควบคุมจำนวนและความหนาแน่นของตัวอักษรที่จุดต่างๆได้

2.3 อุปกรณ์แก้ไขในคอมพิวเตอร์

- กล้องโทรทรรศน์อวกาศ
- เครื่องฉายสไลด์ (Projector)
- คอมพิวเตอร์ (Computer)
- ฟิล์มโค้ด (Gray-Code Pattern)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

3 รูปแบบของภาพและการโฟกัส

(Image Formation and Focusing)

3-4-1

2.1 Image Formation and Focusing

ภาพของวัตถุที่ถูกโฟกัสด้วยเลนส์ (lens) บนอุปกรณ์ตรวจจับ (sensing element) โดยทั่วไป มักใช้กล้องถ่ายรูป (photographic camera) ความแตกต่างระหว่างกล้องถ่ายรูป และระบบการมองเห็นของเครื่องจักร (machine vision) คือ กล้องถ่ายรูปจะใช้ฟิล์ม ในขณะที่ระบบการมองเห็นของเครื่องจักรจะใช้ความคมชัด (contrast) เพื่อที่จะจับ (capture) ภาพ (image) อุปกรณ์ตัวเซ็นเซอร์จะเปลี่ยนภาพที่มองเห็น (visual image) ไปเป็นสัญญาณทางไฟฟ้า (electrical signal)

กล้องที่ใช้ในระบบการมองเห็น (machine vision camera) โดยปกติมีรายละเอียดแตกต่างจากระบบการมองเห็น (vision system) ขณะเดียวกันความสามารถและความต้องการต่างๆของกล้องก็ขึ้นอยู่กับการใช้งาน พารามิเตอร์ของเลนส์กล้องก็ต้องบอกรายละเอียดเอาไว้ เพราะว่ามันเป็นอุปกรณ์ที่จะปรับกล้องให้ได้รายละเอียดตามการใช้งาน พารามิเตอร์ที่สำคัญ 4 ประการที่มีส่วนร่วมกับออปติคอลเลนส์ (optical lens) ของระบบมองเห็น คือ

- 1) Magnification
- 2) Focal length
- 3) depth of field
- 4) Lens mounting

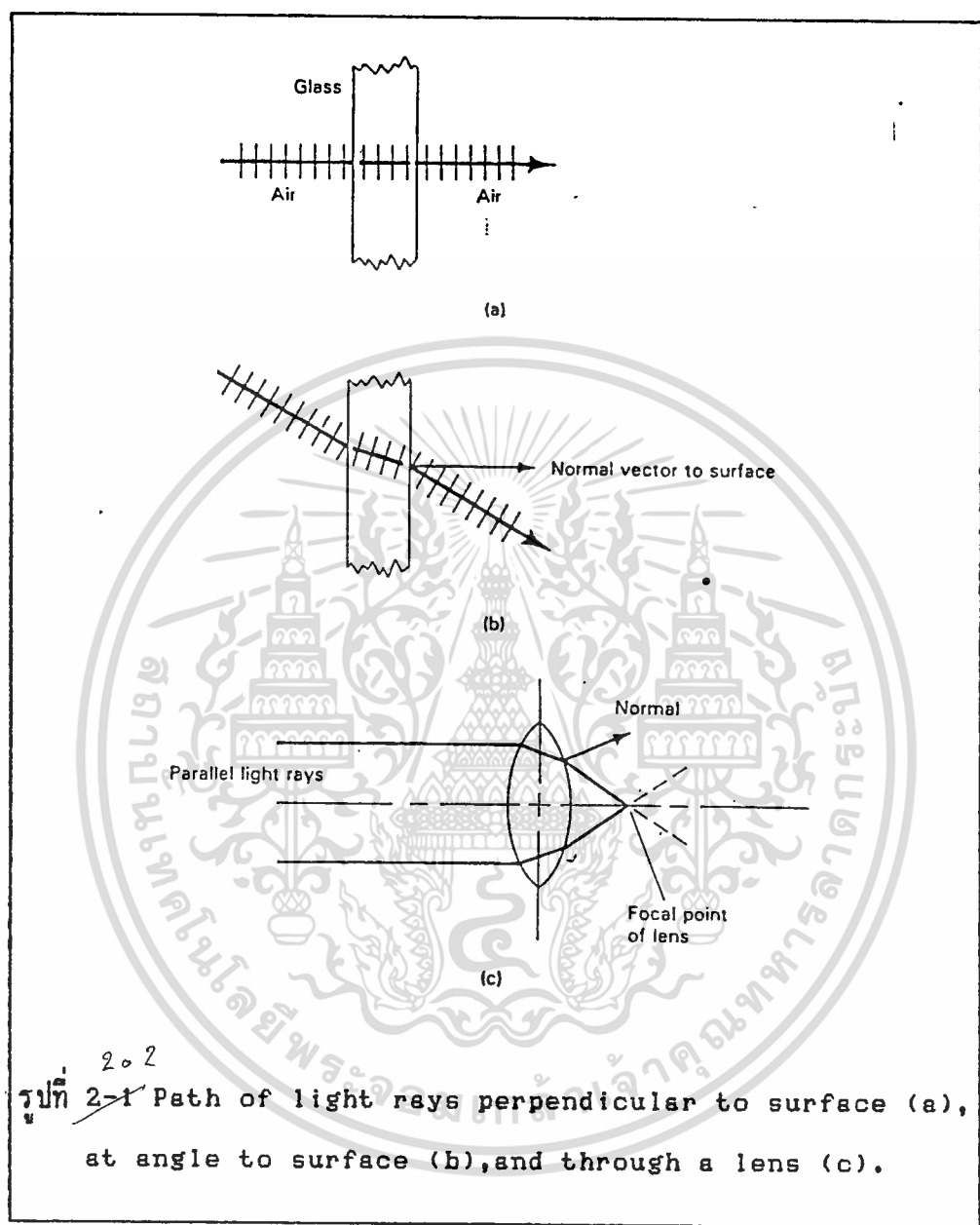
ก่อนที่จะกล่าวถึงพารามิเตอร์ ทั้ง 4 ข้างต้น ขอกล่าวถึงเรื่องเลนส์ไว้สักเล็กน้อย

คุณสมบัติการโฟกัสของเลนส์ เป็นผลจากคลื่นแสง (light wave) ที่มีความเร็วในการเดินทางในอากาศเร็วกว่าในแก้ว หรือ วัสดุพวกออปติคอล (optical) ลำแสงที่ตัดผ่านรอยต่อจากอากาศไปยังแก้ว จะเบี่ยงเบนไปจากเวกเตอร์ปกติ (normal vector) ดังรูป 2-1

ลำแสงที่พุ่งออกจากรอยต่อระหว่างแก้ว-อากาศ จะเบี่ยงเบนไปจากปกติของผิวเชื่อมต่อระหว่าง แก้ว-อากาศ ดังนั้น ลำแสงที่ตั้งฉากรอยต่อระหว่าง อากาศ-แก้ว-อากาศ จะไม่เบี่ยงเบนไปจากเดิม ดังแสดงในรูปที่ 2-1 (a) และมุมของการตกกระทบรอยเชื่อมต่อกับมุมที่ลำแสงผ่านออกมาจากรอยเชื่อมต่อจะเป็นมุมเดียวกัน มันจะถูกแทนที่โดยการรวมของแสงซึ่งขึ้นอยู่กับความกว้างและดัชนีหักเหของแก้ว (refractive index) ถ้าผิวทั้งสองของแก้วขนานกันจะได้ผลดังรูปที่ 2-1 (b)

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในกรณีที่ผิวทั้งสองข้างของเลนส์เป็นผิวโค้ง ลำแสงที่ผ่านไปยังเลนส์จะเปลี่ยนเป็นจุด ซึ่งรู้จักในนาม "จุดโฟกัส" ซึ่งตำแหน่งของจุดโฟกัส สามารถหาได้จากสมการข้างล่าง



$$\frac{1}{f} = \frac{1}{R_1} + \frac{1}{R_2}$$

เมื่อ f คือระยะโฟกัส (focal length), ระยะจากเซ็นเตอร์เพลน (center plane) ของเลนส์

R_1 และ R_2 เป็นรัศมีของส่วนโค้งของผิวเลนส์ทั้งสอง

ลำแสงจากวัตถุ (object) จะเป็นเส้นตรงที่ต่อเนื่องไปยังจุดโฟกัส และรูปแบบของภาพบนตำแหน่งระนาบ f ระยะเกินจุดโฟกัส ภาพจะเป็นภาพกลับหัว และขนาดของภาพจะเพิ่มขึ้น
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขณะที่ระยะจากเลนส์เพิ่มขึ้น อุปกรณ์ตรวจจับ หรือ เซ็นเซอร์อะเรย์ (sensor array) ของระบบมองเห็นอยู่ ณ ตำแหน่งคงที่ในบริเวณของระบบมองเห็นในตัวกล้อง ตำแหน่งของตัวตรวจจับ (detector) จะถูกสร้างขึ้นโดยขึ้นอยู่กับโครงสร้างของกล้องและจะไม่แปรเปลี่ยนองค์ประกอบ (variable factor) ไปในการประยุกต์ใช้งาน

๒. Magnification

อัตราขยาย (Magnification) m คือการวัดความสัมพันธ์ระหว่างขนาดของภาพที่มองเห็น (visual image) ของวัตถุในความจริง (physical world) ต่อขนาดของภาพบนตำแหน่งตัวเซ็นเซอร์ ณ ระนาบตรวจจับ (detector plane) ในตัวกล้อง ค่าของ m จะมีค่าน้อยกว่าหนึ่ง ในกรณีที่ประยุกต์ใช้งานในงานอุตสาหกรรม เพราะว่าขนาดตัวตรวจจับ (detector) มีขนาดเล็กกว่าวัตถุ (object) ที่เรามองเห็น และค่าของ m จะมีค่ามากกว่าหนึ่ง ในกรณีของกล้องจุลทรรศน์

ความสูงของภาพ (image) H_i ในรูปที่ 2-2 เป็นฟังก์ชันของระยะทางระหว่างเลนส์และตำแหน่งของอิมเมจเพลน (image plane) ในตัวกล้อง ถ้าจุดสูงสุดของวัตถุในรูป 2-2 ถูกกำหนดให้เป็น o และจุดสูงสุดของภาพรูปลูกศรที่ i รูปสามเหลี่ยม (o, o, D_o) เป็นสามเหลี่ยมคล้ายกับสามเหลี่ยม (c, i, D_i) และอัตราส่วนของด้านจะเท่ากัน อัตราขยาย (magnification) คือ อัตราส่วนของขนาดภาพ (image size) ต่อขนาดวัตถุ (object size)

$$m = \frac{H_i}{H_o} + \frac{D_i}{D_o}$$

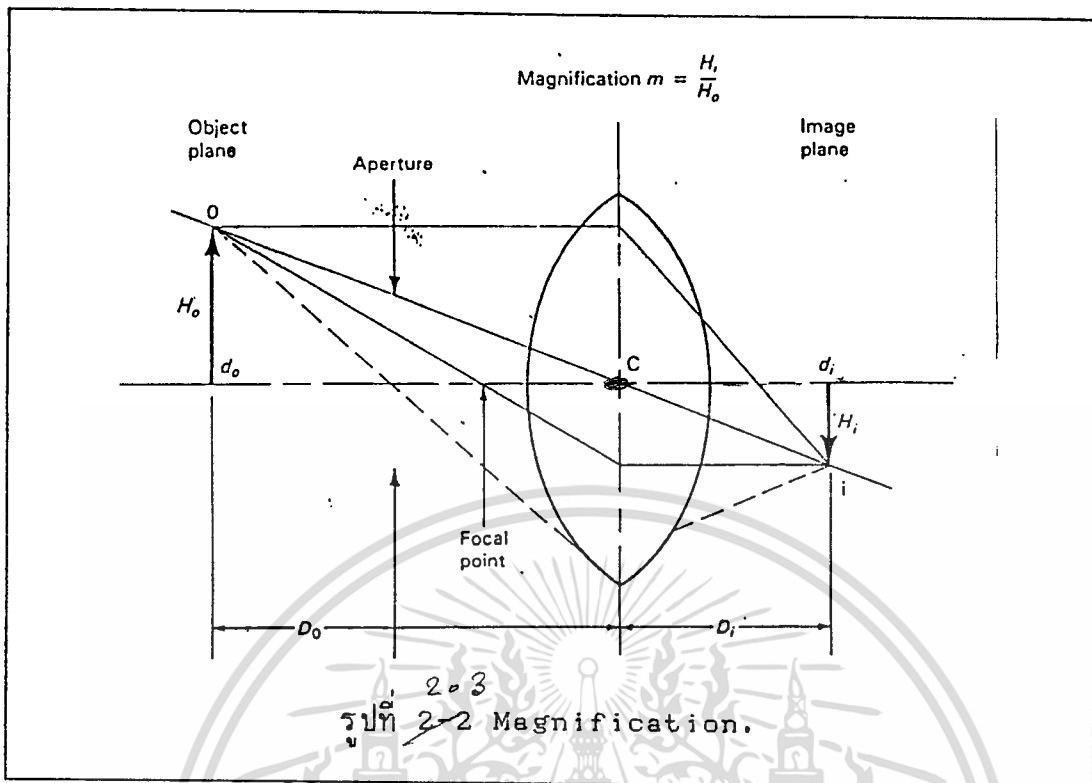
เมื่อ H_o เป็นด้านของเซ็นเซอร์อะเรย์ (sensor array)

D_o เป็นระยะทางระหว่างเลนส์และระนาบของภาพ (image plane)

D_o ถูกตรึง (fixed) ให้คงที่ เมื่อกล้องมีโครงสร้างตามที่ถูกสร้างออกมา D_o เป็นระยะระหว่างวัตถุและเลนส์ของกล้อง

คุณสมบัติของอัตราขยายถูกวางบนพื้นฐานขนาดสูงสุดของ H_o (H_o เป็นขนาดของภาพซึ่งถูกโปรเจกต์ไปยังเซ็นเซอร์อะเรย์ D_o ก็คือ ระยะระหว่างเลนส์และโปรเจกต์เพลน) การเพิ่มขึ้นของระยะ D_o จะทำให้อัตราการขยายลดลง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



$d_i = \frac{d_o}{m}$

ตัวอย่าง

Determine the magnification of the vision system and the size of a pixel on the sensor and on the object if the size of the 100 x 100 solid state sensor array is 0.15 x 0.15 inches and the size of the object to be measured is 3 by 3 inches.

Magnification

The magnification is the ratio of image to the object size.

$d_i = \frac{d_o}{m} \times d_o$

$$m = \frac{H_i}{H_o} = \frac{0.15}{3.0}$$

Pixel size on array

the size of a pixel is equal to the distance between sensor elements. The pixel size is determined by dividing the dimension of the sensor array by the number of elements in the same direction.

$$\text{Pixel size} = \frac{\text{dimension of array}}{\text{number of elements}} = \frac{0.15}{100} = 0.0015$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้เฉพาะในการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Pixel size on object

The size of the pixel on the object can be determined from the system magnification.

$$\text{Pixel size on object} = H_o = \frac{H_i}{m} = \frac{0.0015}{0.05} = 0.03$$

Focal length

คุณลักษณะของการโฟกัส (Focal characteristics) หรือ ระยะโฟกัส (Focal length) f ของเลนส์แบบธรรมดา ถูกกำหนดให้มีความสัมพันธ์กับความโค้งของผิวเลนส์ทั้งสอง ดังสมการข้างล่างนี้



หรือ

$$\frac{1}{f} = \frac{1}{R_1} + \frac{1}{R_2}$$

$$f = \frac{R_1 \times R_2}{R_1 + R_2}$$

เมื่อ R_1 และ R_2 เป็นรัศมีความโค้งของผิวเลนส์ ระยะโฟกัส (Focal length) มีความสัมพันธ์กับพารามิเตอร์ของระบบ 2 ตัว คือ อัตราการขยาย (magnification) และระยะจากเลนส์ถึงวัตถุ ดังสมการ

$$f = \frac{1}{1 + 1/m}$$

การออกแบบของคุณสมบัติระยะโฟกัส (Focal length) จะทำให้เกิดระยะระหว่างและวัตถุทำให้เกิดอัตราขยาย (magnification)

ตัวอย่าง

Determine the focal length of the lens for a vision system with a magnification of 0.05 and a distance of thirty inches between the object and camera.

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิธีทำ

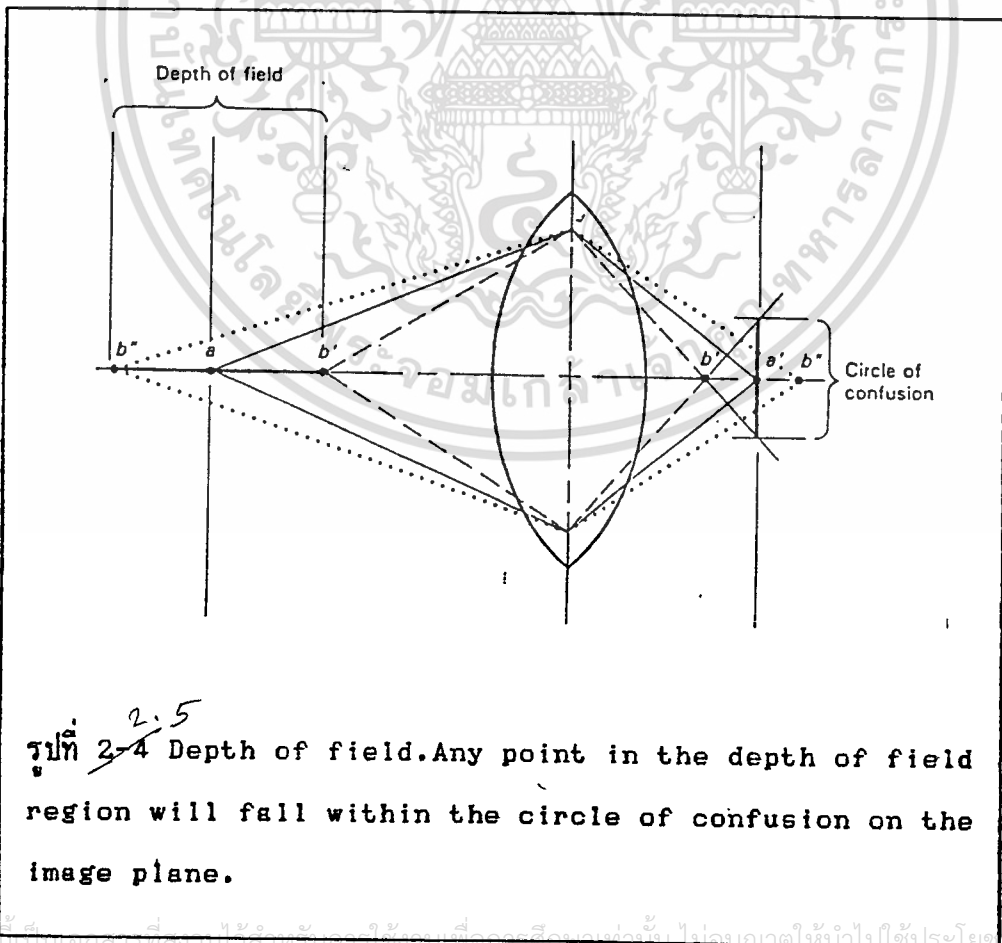
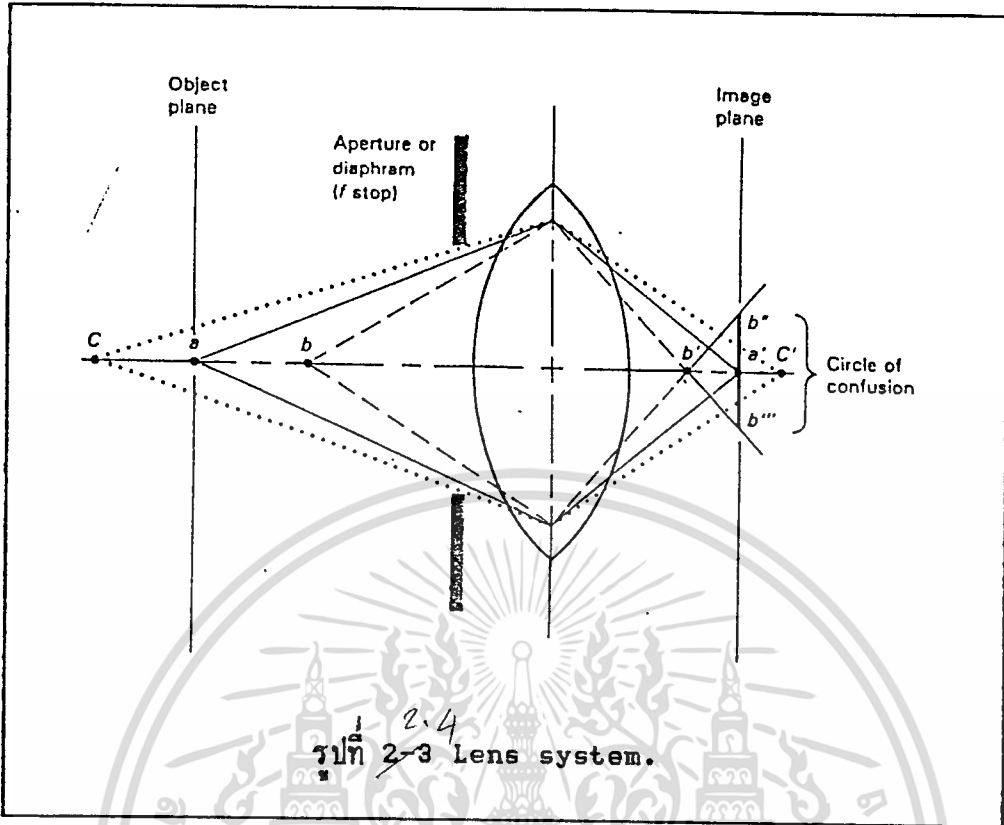
Focal length

$$FL = \frac{D_o}{1 + 1/m} = \frac{30''}{1 + 1/0.05} = 1.428''$$

~~ภาพสองมิติที่อยู่ตำแหน่ง 30'' จากเลนส์จะเป็นตำแหน่งโฟกัสที่ดีที่ลดบนอิมเมจเพลน + มี~~
 เลนส์อยู่ในสภาพดี ในการใช้งานในงานอุตสาหกรรม มักจะเกี่ยวกับวัตถุที่เป็นสามมิติ ผลกระทบ
 ของจุดที่พิจารณาวัตถุที่ไม่อยู่บนออบเจกต์เพลน (object plane) จะทำให้ภาพ (image)
 ไม่ชัด (blur) นั่นคือ จุดบนวัตถุจะหนีบ (map) ไปเป็นภาพสองมิติบนอิมเมจเพลนดังแสดงในรูป
 ที่ 2-3 ขณะที่ลำแสงลู่ออกจากจุด ณ ระยะที่แตกต่างจากระยะโฟกัส ความสัมพันธ์ระหว่าง
 circle of confusion และ depth of field ถูกแสดงไว้ดังรูปที่ 2-4

จุด a บนออบเจกต์เพลน (object plane) ที่หนีบ (map) ไปบนจุด a' บนระนาบ
 ของภาพ (image plane) จุด b อยู่ระหว่างออบเจกต์เพลนและเลนส์ จะหนีบไปบนจุด b' ที่
 ตำแหน่งระหว่างอิมเมจเพลนและเลนส์ มันจะหนีบไปบนพื้นที่ b''-b' ที่รู้จักกันในนาม
 "circle of confusion" บนอิมเมจเพลน ในทำนองเดียวกันจุด c ที่ระยะมากกว่าออบเจกต์
 เพลนจะหนีบ (map) ไปบนจุด c' ณ ระยะที่มากกว่าระยะของอิมเมจเพลนไปยังเลนส์ อิมเมจ
 จะไม่ชัด หรือไม่อยู่ในตำแหน่งโฟกัส (out of focus) ในกล้องถ่ายภาพที่นิยมใช้กัน

depth of field คือ บริเวณซึ่งไม่ได้โฟกัสจะไม่มีผลต่อคุณภาพของข้อมูลที่วัดได้ ทุกจุดที่
 พิจารณาวัตถุสามมิติอยู่บนพื้นผิวที่การตรวจจับ (inspection) จะต้องอยู่ใน depth-of-
 field ดังแสดงในรูป 2-5 ขนาดของ D_o คิดตั้งแต่จุดศูนย์กลางของ depth-of-field โซน
 (zone) ไปยังเลนส์



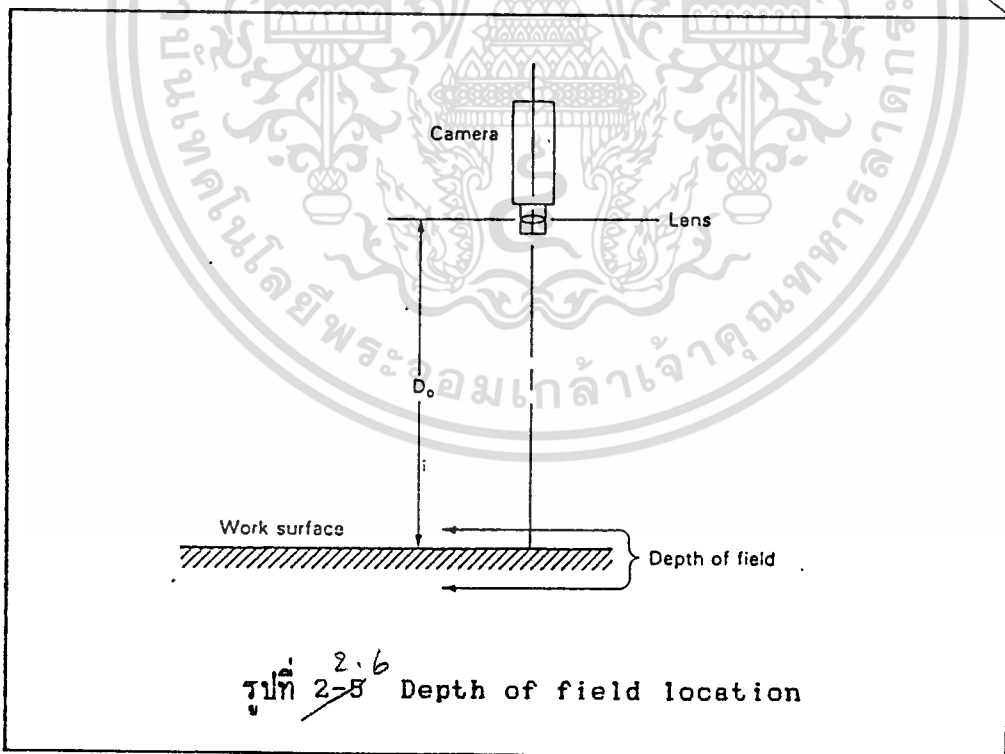
เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ในเชิงพาณิชย์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Depth of field

ช่องว่างที่อยู่เหนือและใต้ของออบเจกต์เพลน ซึ่งเลนส์ยังรักษาการโฟกัสของอิมเมจให้สามารถรับชัดจำกัดใน depth of field ซึ่งเป็นเทคโนโลยีเกี่ยวกับการมองเห็น เมื่อมองวัตถุอยู่บนผิวชิ้นงาน (work surface) ซึ่งเป็นจุดที่พิจารณาไม่อยู่ที่ออบเจกต์เพลน depth of field เป็นฟังก์ชันของขนาดช่องรับแสงในกล้อง อัตราการขยาย (magnification) และขนาดของตัวตรวจจับ (sensor elements) ความลึก (depth) เพิ่มขึ้นขณะที่ช่องรับแสง (aperture) เล็กลง แต่การรวมของแสงส่งไปกลับลดลง การปรับช่องรับแสง (aperture) ของกล้อง ใช้การเปลี่ยนขนาดของการเปิดหน้ากล้องของเลนส์ ที่แสดงไว้ในรูป 2-3 f มีค่าเท่ากับ 16 หมายถึงการเปิดช่องรับแสง (aperture) เล็กที่สุด การเปิดช่องรับแสงเพิ่มขึ้น ค่า f จะลดลง

ตัวเลขที่ใกล้เคียงกันบนการลดลงตามลำดับของสเกลที่เพิ่มขึ้น ความสัมพันธ์ การรวมแสง การส่งโดยองค์ประกอบสองตัว ค่า f มาตรฐาน คือ 16, 11, 8, 5.6, 4, 2.8, 2 และ 1.5

ขนาดของช่องรับแสงมีผลต่อระบบการมองเห็นในการเปิดช่องรับแสงของเลนส์ที่เล็กอยู่สองประการ คือ



- 1) Depth of field มากขึ้น
- 2) การรวมของแสงน้อยลงที่จะส่งไปยังตัวเซ็นเซอร์ (sensor)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ค่า f-stop ที่มากจะเพิ่มย่านการทำงาน (working range) ให้กว้างขึ้น แต่ต้องาร เวลาการเปิดหน้ากล้องมากขึ้น เวลาเปิดหน้ากล้องที่นาน จะไม่ทำให้เกิดปัญหาเรื่องโพสท่าสำหรับภาพนิ่ง (static image) แต่มีผลกระทบต่อความสามารถในการวัด ภายในสภาวะไดนามิก (Dynamic condition)

พื้นที่พิกเซลของตัวเซ็นเซอร์เล็กๆ จะให้ค่า depth of field น้อยลง ความสัมพันธ์ของ depth of field และพารามิเตอร์ของกล้อง คือ

$$\text{Depth of field} = \frac{2af(m+1)}{m^2}$$

เมื่อ a คือ ขนาดของพิกเซล

f คือ ขนาดของการเปิดหน้ากล้อง (f-stop of focus)

m คือ อัตราการขยาย (magnification)

ตัวอย่าง

Determine the depth of field for a vision system having a 200 x 200 array sensor 0.30 x 0.30 inches, an f-stop of 16, and magnification factor of 0.05.

$$\text{Pixel size} = 0.30/200 = 0.0015''$$

$$2 \times 0.0015 \times 16 \times (1+0.05)$$

$$\text{Depth of field} = \frac{\quad}{(0.05)(0.05)} = 20.16''$$

$$(0.05)(0.05)$$

ถ้าความสูงของวัตถุน้อยกว่าครึ่งหนึ่งของ depth of field ระยะทางระหว่างเลนส์และผิวชิ้นงานสามารถที่จะใช้การหาโฟคัลเล็ง (Focal length) ของเลนส์ ในการใช้งานจะมืองค์ประกอบ (factor) ตัวหนึ่ง คือ magnification ซึ่งมีค่ามากกว่าหนึ่ง ค่า depth of field ที่สัมพันธ์กันจะเล็ก ค่า depth of field ที่ใช้งานแสดงให้เห็นในรูป 2-6 ขีดจำกัด

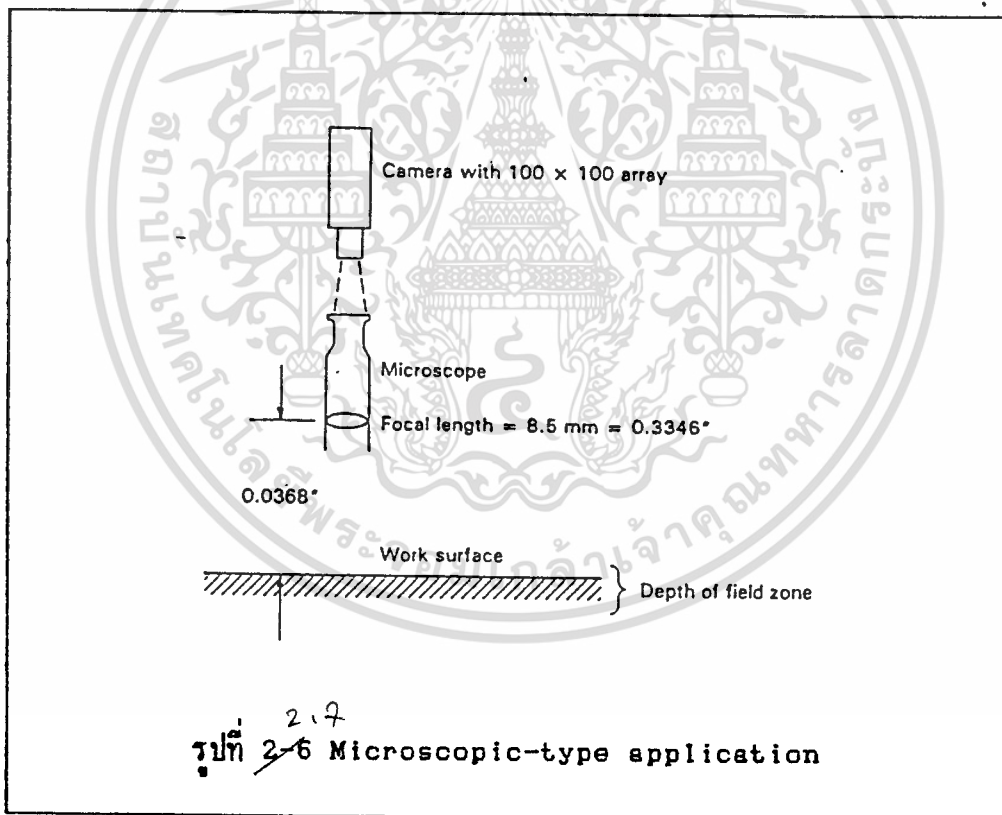
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขนาดของผิวขรุขระ ซึ่งสามารถมองเห็นได้

ระยะทางระหว่างเลนส์และผิวชิ้นงาน คือ 0.0368" กับค่า f -stop เซ็ตไว้ที่ค่า 16 คำนวณหา depth of field สูงสุด ขนาดพิกเซลบน 100 x 100 , 0.15" อะเรย์เซ็นเซอร์ เป็น 0.0015" อัตราขยาย (magnification) คือ 10

$$\begin{aligned} \text{depth of field} &= 2 \times 0.0015 \times 16 \times (1+0.05)/(10 \times 10) \\ &= 0.00528" \end{aligned}$$

ผิวที่ขรุขระเท่ากับครึ่งหนึ่งของ depth of field คือ 0.00264" สามารถเป็นอิมเมจได้ ถ้าผิวของวัตถุอยู่ที่อิมเมจเพลน



Lens mountings

ชนิดของเลนส์เมอตัง (lens mountings) จำต้องมีลักษณะเฉพาะ เมื่อซ็อกกล้องและเลนส์ที่ใช้งานกันอยู่ มีด้วยกัน 4 ชนิด คือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- 1) C-mounting for machine vision
- 2) U-mounting for 35 mm camera
- 3) L-mounting for fixed flat field installation and
- 4) Bayonet for quick change 35 mm camera

เลนส์เมาต์ติ้ง (lens mounting) ที่ใช้ในปัจุบัน คือ แบบ C-mount ตัวกล้องกำลังเปลี่ยนมาใช้แบบใช้เมาท์ (mount) ชนิดสลักบนตัวกล้อง และจัดเตรียมวงแหวนอะแด็ปเตอร์ (Adapter) ให้เข้ากับ C-mount lens ได้

ระยะระหว่าง flange ของ C-mount และ อิมเมจเพลน (image plane) ระยะของจุดโฟกัส คือ 17.5 mm หรือ 0.69" เส้นผ่านศูนย์กลางของพื้นที่ที่สามารถใช้ได้ คือ ตั้งแต่ขนาดของอะเรย์จนถึง 0.512" สามารถที่จะใช้ได้ จะมีสกรูยึดกับเกลียวขนาด 1" 23 เกลียวต่อไปนี้

ส่วนแบบอื่นๆนอกจากนี้จะไม่ขอกล่าวถึง

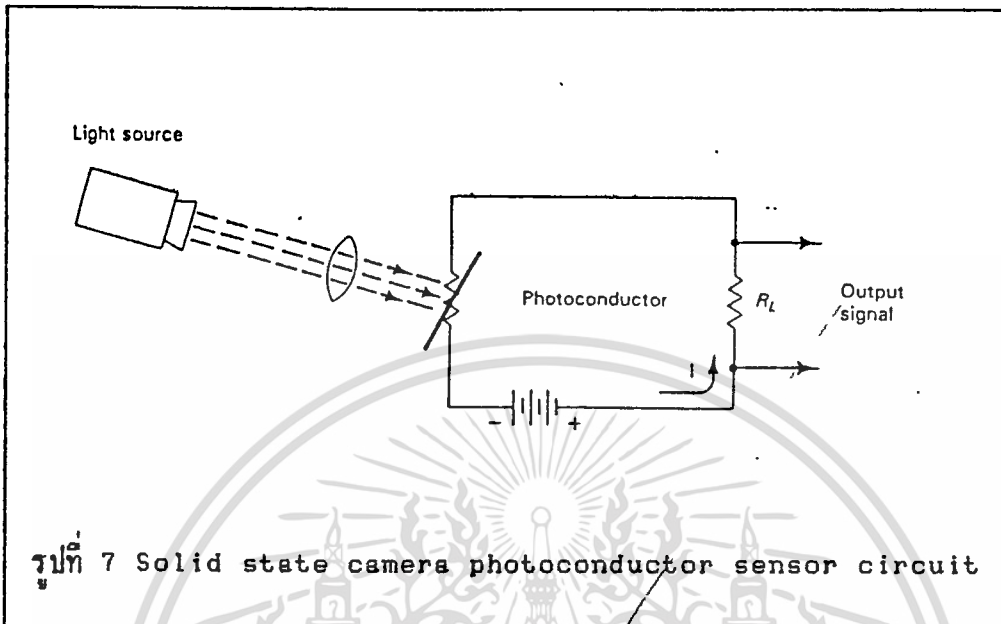
Solid state camera

หลักพื้นฐานของ Solid State Image Sensors เป็นการแยกสัญญาณไฟฟ้าที่ถูกสร้างมาจากแต่ละพิกเซล หรือ พื้นที่ในตัวเซ็นเซอร์ตัวนี้จะบรรลู่ได้โดยการวางตำแหน่งของโฟโตคอนดักเตอร์อิลีเมนต์ (photoconductor element) ซึ่งมีคุณสมบัติในการลดความต้านทาน เมื่อมีการรวมแสงมาตกกระทบบนอุปกรณ์ ในแต่ละพื้นที่ของพิกเซล สัญญาณเอาท์พุทโวลเตจ ดังแสดงในรูป 2-7 คือขึ้นอยู่กับความต้านทานของโฟโตคอนดักเตอร์

ความสัมพันธ์ระหว่างการเปลี่ยนพลังงานแสง และเอาท์พุทโวลเตจไม่เป็นเชิงเส้น และขึ้นอยู่กับ transfer characteristics ของโฟโตคอนดักเตอร์ อัตราส่วนการเพิ่มขึ้นของกระแสไฟฟ้าต่อการเพิ่มขึ้นของพลังงานแสง ถูกอ้างอิงได้โดย Gramma Power

เอาท์พุทของวงจรจะเป็น voltage signal ที่สอดคล้องกับค่าความเข้มของแสงบนพิกเซลที่มีความไวโดยอุปกรณ์สารกึ่งตัวนำ (semiconductor element) ตัวตรวจจับ (detector) จะใช้กันหลายๆตัว ถูกจัดเรียงให้เป็นเชิงเส้น และอยู่ในกรอบสี่เหลี่ยมของอะเรย์และข้อมูลภาพ (image data) หาได้จากการสุ่ม (sampling) เอาท์พุทของตัวดีเทคเตอร์แต่ละตัว ตัวดีเทคเตอร์จะต้องมีความไวในการตรวจจับแสงที่ตกกระทบ และมีค่าเวลาคงที่ (time constant) ซึ่งยอมให้การสุ่ม (sampling) ของอิลีเมนต์ทั้งหมดในอะเรย์ เพื่อหาข้อมูลภาพ (image data) เอาท์พุทของตัวเซ็นเซอร์เป็นอนุกรมของโวลเตจพัลส์ที่แทนความเข้มของแสงที่ตำแหน่งของพิกเซล โวลเตจพัลส์จะต้องเป็นดิจิทัล Solid state camera จะไม่มีปัญหาใน

เรื่อง blooming และ flare ความเข้มของแสงที่ชิป (chip) สามารถมองได้โดยตรงกับความน่าจะเป็นที่จะทำให้ชิป (chip) เกิดความเสียหายได้



วัสดุที่ใช้ทำโซลิตสเทท (solid state) มีด้วยกัน 4 ประเภทหลักๆ คือ

- 1) MOS - Metal Oxide Semiconductor
- 2) CCD - Charge Coupled Device
- 3) CID - Charge Injected Device
- 4) CPU - Charge Priming Device

TABLE 2-1 Solid State Camera Performance

Chip Type	MOS	CCD	CPD/CID
Sensitivity (minimum lux)	10	3	5
Dynamic Range (relative)	1	4	1
	(poorest)		
Dark Current	good	good	good
Noise	15	1	4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับนอร์ใช้เฉพาะเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

CID และ CPD จะมีคุณลักษณะที่เหมือนกันจึงจับเป็นกลุ่มเดียวกันในการเปรียบเทียบ ดังใน ตาราง 2-1

คุณลักษณะของนิโคเซลทั้งแถวจะไม่ทำงาน ถ้านิโคเซลตัวใดตัวหนึ่งเกิดเสียหายในกล้องแบบ CCD ทำให้กล้องประเภทนี้ไม่นิยมใช้ในงานอุตสาหกรรม นอกจากนั้น ส่วน CCD (CCD unit) ยังเป็น moire pattern ซึ่งจะปรากฏเป็นวงแหวนซ้อนๆกันบนหน้าจอมอนิเตอร์ ๗

ขนาดของ integrated circuit array จะมี 256 elements ในหนึ่งแถว จะน้อยกว่า 1/4 นิ้ว และบรรจุ ground-polished window สำหรับป้องกันความเสียหายที่จะเกิดขึ้นกับ photodiode element

Solid state photodiode array สามารถที่จะประดิษฐ์ให้มันเชิงเส้น (linear) หรือเป็น ระเบียบแบบสี่เหลี่ยมผืนผ้า (rectangular array) ระเบียบที่เชิงเส้นสามารถที่จะประดิษฐ์ในรูปของวงกลมที่มีช่องห่างเท่าๆกัน สำหรับใช้ในการโฟกัส แตกคิง (tracking) และ มุมที่แตกต่างออกไป วงกลมหนึ่งอันกับ 720 elements จะให้ 0.5 องศาของช่องว่าง

Solid state camera มีน้ำหนักเบาเมื่อเปรียบเทียบกับชนิดหลอด จะประมาณ 1 ใน 15 ส่วนของน้ำหนักกล้องชนิดหลอด ใช้พลังงานเพียง 2 Watt และใช้กับไฟ battery 12 volts และมีอายุการใช้งานประมาณ 7-10 ปี

Characteristic of solid state camera

Solid state camera สามารถบรรจุเส้นสแกนได้ที่ละเส้น หรือเป็น matrix array sensors ตัวอักษรจะประกอบด้วยตัวรับแสง (light-sensing element) หลายๆแถวซึ่งจะสร้างเป็นอิมเมจสองมิติ ๗ ต่อ

คุณลักษณะของ Solid state มีส่วนต่างๆดังนี้

- 1) Matrix array cameras are available with RS-170 format output - for display on a TV monitor.
- 2) Line scan units can be attached end to end to produce longer - linear arrays; the matrix array size range from three to five hundred by three to five hundred pixels and are increasing with time.
- 3) Response is highest in the infrared region of the spectrum and may require an infrared correction filter for normal image.

๔) Cameras and sensing arrays are available with different levels

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- of bad or malfunctioning pixels depending on the price.
- 5) There is virtually no geometric distortion, drift, or lag.
 - 6) Cameras are light and rugged and consume little power ; they can be operated from batteries.
 - 7) Pixel transfer rates can be as high as 20 megapixels per second.
 - 8) They can have C,U, or bayonet lens mounts with adapter ring for a C-mount lens.

2.4.2 2.2 Processing

ขีดจำกัดทางกายภาพของฮาร์ดแวร์ที่ใช้ในการรับเอาอิมเมจเริ่มต้นจะมีทั้ง random และ systematic noise ในข้อมูลภาพ (image data) ดังนั้นอิมเมจอาจบรรจุข้อมูลบดบังลักษณะที่เด่นซึ่งไม่ได้กำลังพิจารณาอยู่ ฟังก์ชันพื้นฐานของการประมวลผลภาพ (image processing) เป็นการสร้างอิมเมจใหม่ โดยการเลือกข้อมูลที่เป็นจุดเด่นที่น่าสนใจ จะถูกทำให้ดีขึ้นและผลกระทบจากสัญญาณรบกวน (noise) ก็จะถูกลดลงหรือกำจัดออกไปเลย

Processing element จะมีวงจรอิเล็กทรอนิกส์และโปรแกรมซอฟต์แวร์ที่จะต้องปฏิบัติ

- 1) Image grabbing
- 2) Image enhancement
- 3) Feature extraction
- 4) Output formatting

Data processing ในรายละเอียดของระบบการมองเห็นอาจจะถูกทำโดยฮาร์ดแวร์หรือซอฟต์แวร์ โดยทั่วไปการใช้ฮาร์ดแวร์ในการทำงานจะให้ผลที่เร็วต่อระบบ การเลือก data processing algorithm จะเป็นขั้นตอนสุดท้าย ฮาร์ดแวร์ที่เพิ่มมากขึ้น จะทำให้ต้นทุนสูงขึ้น ลดความอ่อนตัวในการประมวลผล และการเพิ่มขึ้นของฮาร์ดแวร์อันนี้ มีความเป็นไปได้ที่ระบบทั้งหมดจะมีลักษณะที่คล้ายกัน และมีผลการประมวลผลที่เหมือนกัน

ส่วนฮาร์ดแวร์ (hardware element) รวมทั้ง signal digitizer และ frame grabber, storage memory ทั้งสองส่วนนี้ถูกบรรจุไว้ในทุกๆระบบ มันเป็นสิ่งจำเป็นที่ใช้ในการเปรียบเทียบถึงฮาร์ดแวร์และวัตถุประสงค์ของตัวแปรเชิงพาณิชย์ ใน vision processing software program จะ compatible ขณะที่มัลักษณะการทำงานที่ไม่เป็นมาตรฐาน

Processing algorithm พื้นฐานสำหรับเทคนิคต่างๆไป เช่น การกำจัดสัญญาณรบกวน
 เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อใช้ในการศึกษาเท่านั้น ไม่ควรนำไปใช้
 ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(noise elemination) การแต่งขอบภาพให้ดีขึ้น (edge enhancement) การกรอง (filter) และ gray scale modification จะถูกบรรจุในตัวแปรเชิงพาณิชย์ของระบบการมองเห็น (vision system) มันไม่จำเป็นที่จะต้องเขียนซอฟต์แวร์ให้บรรลุตามขั้นตอนพื้นฐานนี้

Mode พื้นฐาน 3 mode ของการเลือกข้อมูล คือ

- 1) Point by point in one image
- 2) Using corresponding point on difference image
- 3) Using regional point in one image

Point by point mode จะ map แต่ละพิกเซลของภาพต้นแบบ (original image) ไปยังอิมเมจใหม่ ซึ่งค่าของพิกเซลแต่ละพิกเซลในอิมเมจใหม่ จะมีความสัมพันธ์กับค่าของพิกเซลที่สอดคล้องกันในอิมเมจต้นแบบ ตัวอย่างอันหนึ่งก็คือ การกลับภาพไบนารีอิมเมจ ซึ่งก็คือการเปลี่ยนค่าพิกเซลที่เป็น "0" ให้เป็น "1" และเปลี่ยนค่าพิกเซลที่เป็น "1" ให้เป็น "0"

Corresponding-point mode จะสร้างภาพใหม่โดยการ correlation แต่ละพิกเซลไปเป็นค่าพิกเซลที่สอดคล้องกันในสองหรือมากกว่าจากภาพต้น (source image) ค่าของมันสามารถรวมได้หลายวิธี ดังนั้นค่าพิกเซลของภาพเมื่องสองภาพ ที่เก็บภาพมาคนละเวลา สามารถที่จะนำมาลบกันได้ เพื่อหาการเปลี่ยนแปลงที่เกิดขึ้นระหว่างคาบเวลาที่ถูกแทรก ข้อมูลภาพสามารถปรับปรุงสำหรับการ calibration ตามประสงค์ ซึ่งการเปลี่ยนแปลงจะกระทำต่อตำแหน่งจุดศูนย์กลาง อาจจะไม่ต้องการความถูกต้อง พื้นที่ข้างเคียงอาจจะเพิ่มขึ้น 7% mode นี้สามารถนำไปใช้รวมข้อมูลจากจำนวนต่างๆของตัวเซ็นเซอร์จากบริเวณของ map บนข้อมูลรวม (composite data) เซ็นเซอร์อาจจะเป็นกล้องถ่ายภาพเทียที่ถ่ายภาพมาในตำแหน่งที่ต่างๆกันของ spectrum เช่น infrared, ultraviolet และ visible

Regional-point mode เกี่ยวกับการคำนวณของค่าตำแหน่งพิกเซลในอิมเมจใหม่บนค่าของจำนวนพิกเซลในบริเวณที่อยู่ข้างเคียงจุดศูนย์กลาง mode นี้สามารถใช้การเฉลี่ยข้อมูลจากพิกเซลในบริเวณข้างเคียง ไปยังจุดตาหนิ (flawed point) ซึ่งเป็นข้อมูลน้อยที่สุดของอิลลิเมนต์ค่าของพิกเซลในอิมเมจใหม่จะเท่ากับค่าเฉลี่ยของ 9 พิกเซล ในบริเวณข้างเคียงที่สอดคล้องกันของพิกเซล เทคนิคการทำให้ภาพดีขึ้น (image enhancement techniques) ทั้งหมดจะใช้ Regional-point processing mode

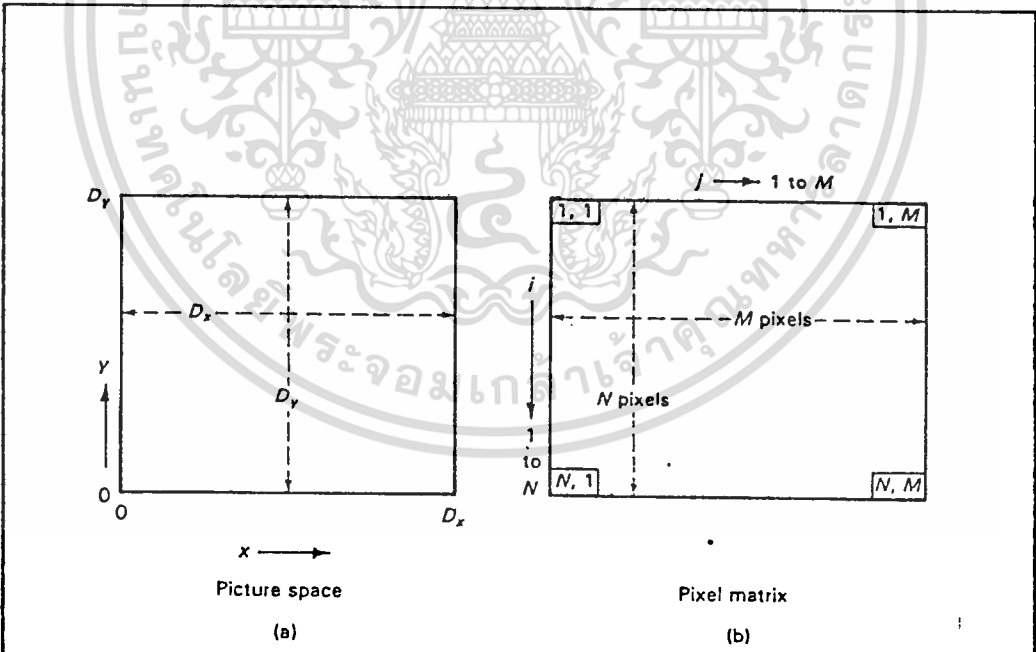
คณ
 หน้า ๓
 บทที่ 3
 การประมวลผลภาพเบื้องต้น

(Fundamental Concepts of Image Processing)

3.1 Pixel

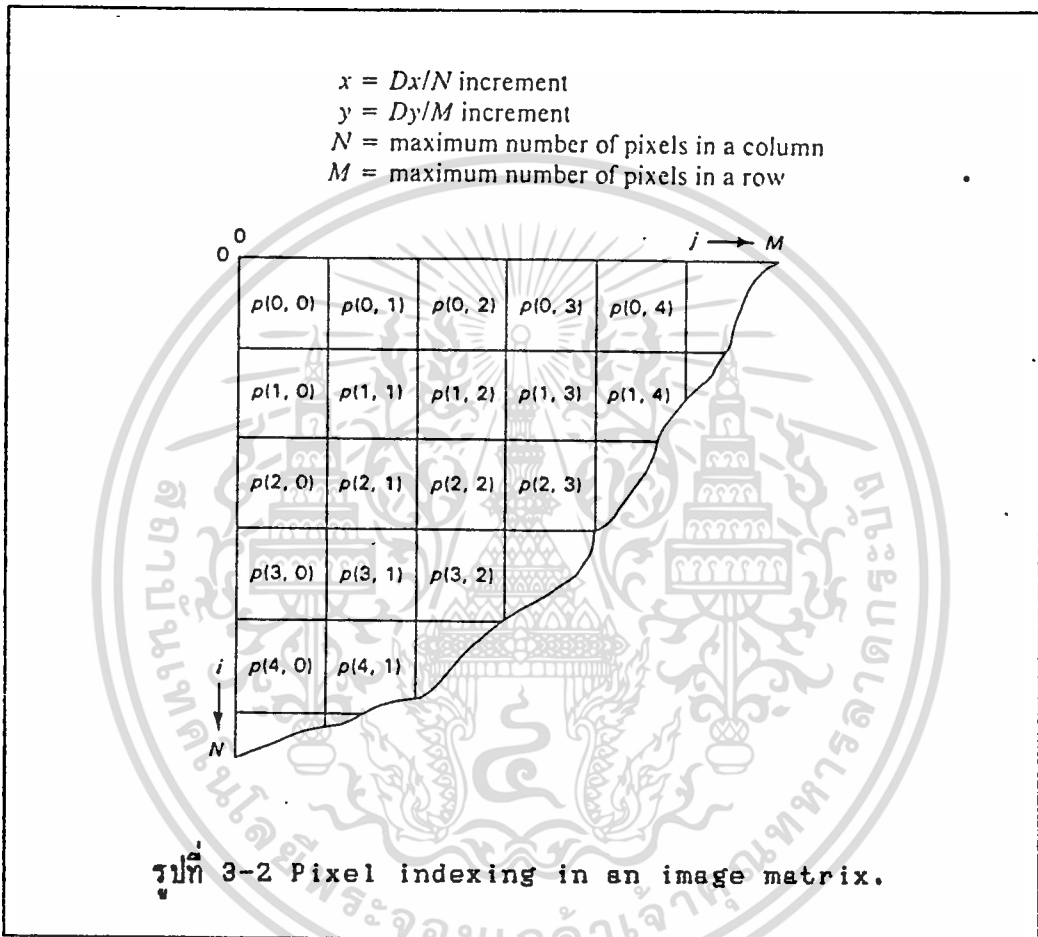
อิมเมจสามารถอธิบายได้โดยค่าของพิกเซลที่มีรูปเป็นเมทริกซ์ (matrix) ขนาด $N \times M$ (ค่าอิลิเมนต์ $P(i, j)$ มีค่าเป็นบวก) ซึ่งแสดงถึงความเข้มของแสงของฟลักซ์ (Flux) บนรูปภาพ อิลิเมนต์ที่ (x, y) จะแทนโดยพิกเซลดังแสดงในรูปที่ 3-1 ซึ่งความสัมพันธ์ระหว่างพิกเจอร์อิลิเมนต์ (picture element) และ พิกเซลเมทริกซ์ (pixel matrix) จุดกำเนิดของภาพ (picture) และ เมทริกซ์ (matrix) จะแตกต่างกัน x และ y เป็นคู่ลำดับ (coordinate) ในภาพ (picture) ที่มีจุดเริ่มต้นทางมุมด้านล่างซ้าย ขณะที่เมทริกซ์ (matrix) มีจำนวนพิกเซลเริ่มต้นที่มุมบนด้านซ้าย

เมื่อ $i = x$ เมื่อ $1 \leq i \leq N$
 $j = (M - y + 1)$ $1 \leq j \leq M$



รูปที่ 3-1 Relationship of picture elements (a) and corresponding pixel distribution in matrix (b).

ค่าจำนวน (numerical value) หรือขนาด (magnitude) ของพิกเซลแสดงให้ด้วยค่าเฉลี่ยความเข้มแสงบนพื้นที่พิกเซลอิเลเมนต์ (picture element) ที่แทนด้วยพิกเซล ค่าของพิกเซล $P_{i,j}$ มีช่วงที่ 0 ถึง 1

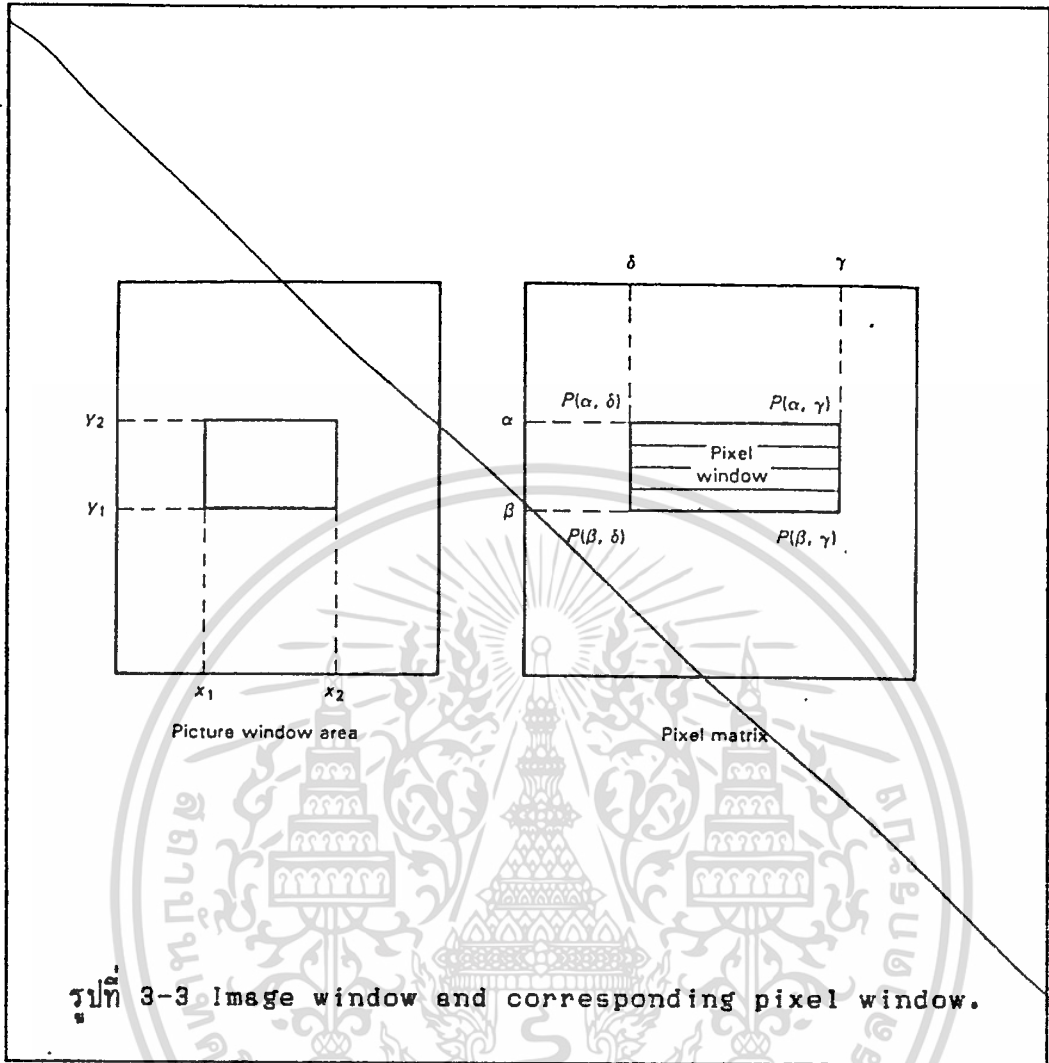


๓๖๕

3.2 Window

บริเวณย่อย (Subregion) ของภาพมักเรียกว่า "window" และถูกกำหนดโดยมุม 4 มุม และค่าของตำแหน่งพิกเซล $P(\beta, \delta), P(\beta, \delta'), P(\alpha, \delta), P(\alpha, \delta')$ ดังแสดงในรูป 3-3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



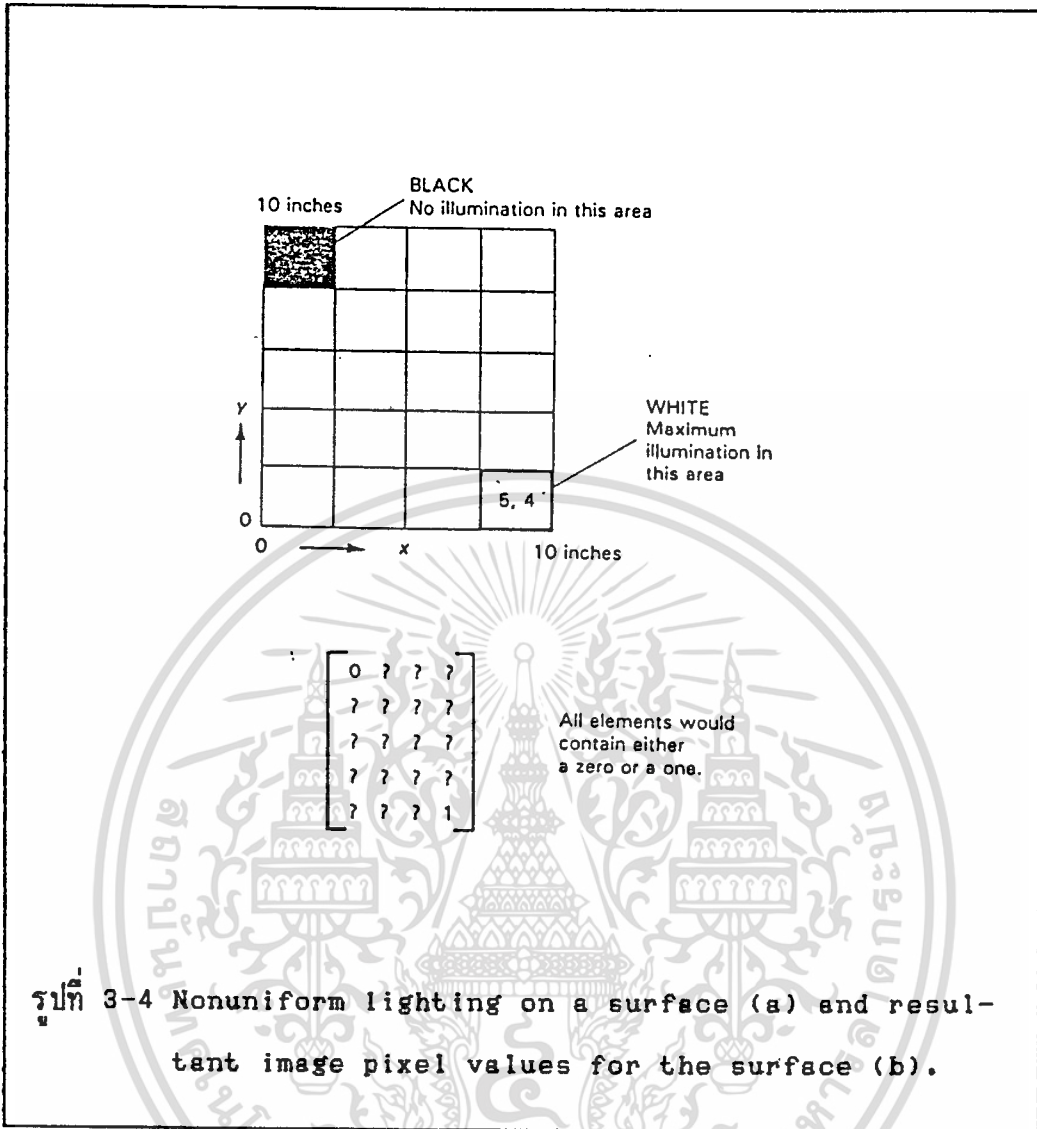
รูปที่ 3-3 Image window and corresponding pixel window.

3.3 Pixel location

คุณสมบัติของพิกเซลในรูปแบบพื้นฐานจะเหมือนกันหมด โดยค่าลำดับ (coordinate) ของพิกเซล จะอยู่ในออร์เรย์ขนาด $N \times M$ ซึ่งแทนเป็นรูปภาพ พิกเซลที่ตำแหน่ง (n, m) จะมีค่าอยู่หนึ่งค่า ซึ่งถูกแทนด้วยค่าเฉลี่ยของแสงที่ตกกระทบลงบนพื้นที่ของภาพ (picture) แทนด้วยพิกเซล

ลองพิจารณาที่ ที่ซึ่งไม่มีแสงตกกระทบบนบริเวณส่วนบนของภาพ และมีแสงที่มีความสว่างมากตกกระทบบนบริเวณส่วนล่างขนาด 10×10 ของภาพ (picture) ดังรูป 3-4 ในระบบของไบนารี จะใช้การแทนค่าของแสงสว่างดังนี้ : บริเวณที่ไม่มีแสงสว่าง จะแทนด้วยค่า "0" และบริเวณที่มีแสงสว่างจะแทนด้วย "1" ภาพจะแสดงโดยเมทริก (matrix) ขนาด 5×4 คือ มี 5 แถว 4 หลัก แต่ละส่วนของภาพ (picture element) มีความกว้าง 2.5 " และ สูง 2 " จะถูกกำหนดเป็นตัวเลข ซึ่งขึ้นอยู่กับค่าเฉลี่ยของแสงที่ตกกระทบบนพื้นที่ส่วนนั้น

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



พื้นที่ขนาด 2.5" x 2" ที่อยู่ส่วนบนทางมุมซ้ายมือของภาพ จะแทนด้วยตำแหน่ง (1,1) ในเมทริก (matrix) ขนาด 5 x 4 จะมีค่า "0" อยู่ แสดงว่าบริเวณนี้ไม่มีแสงตกกระทบ พื้นที่ขนาด 2.5" x 2" ที่อยู่มุมซ้ายส่วนล่างของภาพ ซึ่งอยู่ที่หลักที่ 4 แถวที่ 5 แทนโดยตำแหน่ง (5,4) จะมีค่าเป็น "1" อยู่แสดงว่าบริเวณนี้มีแสงสว่างสูงสุด จุดสังเกตค่าของพิกเซลที่ตำแหน่ง (1,1) จะเป็น "0" และค่าของพิกเซลที่ตำแหน่ง (5,4) จะเป็น "15" ถ้ามีระดับสีเทา (gray scale) เท่ากับ 16 ระดับ ซึ่งนำมาใช้แทนระบบไบนารี

ไม่มีข้อมูลในค่าของพิกเซลที่เป็นค่ากลางๆ และระบบที่ออกแบบไว้จะต้องแสดงค่า "Threshold" (Threshold value) ของแสงสว่าง ซึ่งค่าจะถูกเปลี่ยนจาก 0 ไป 1

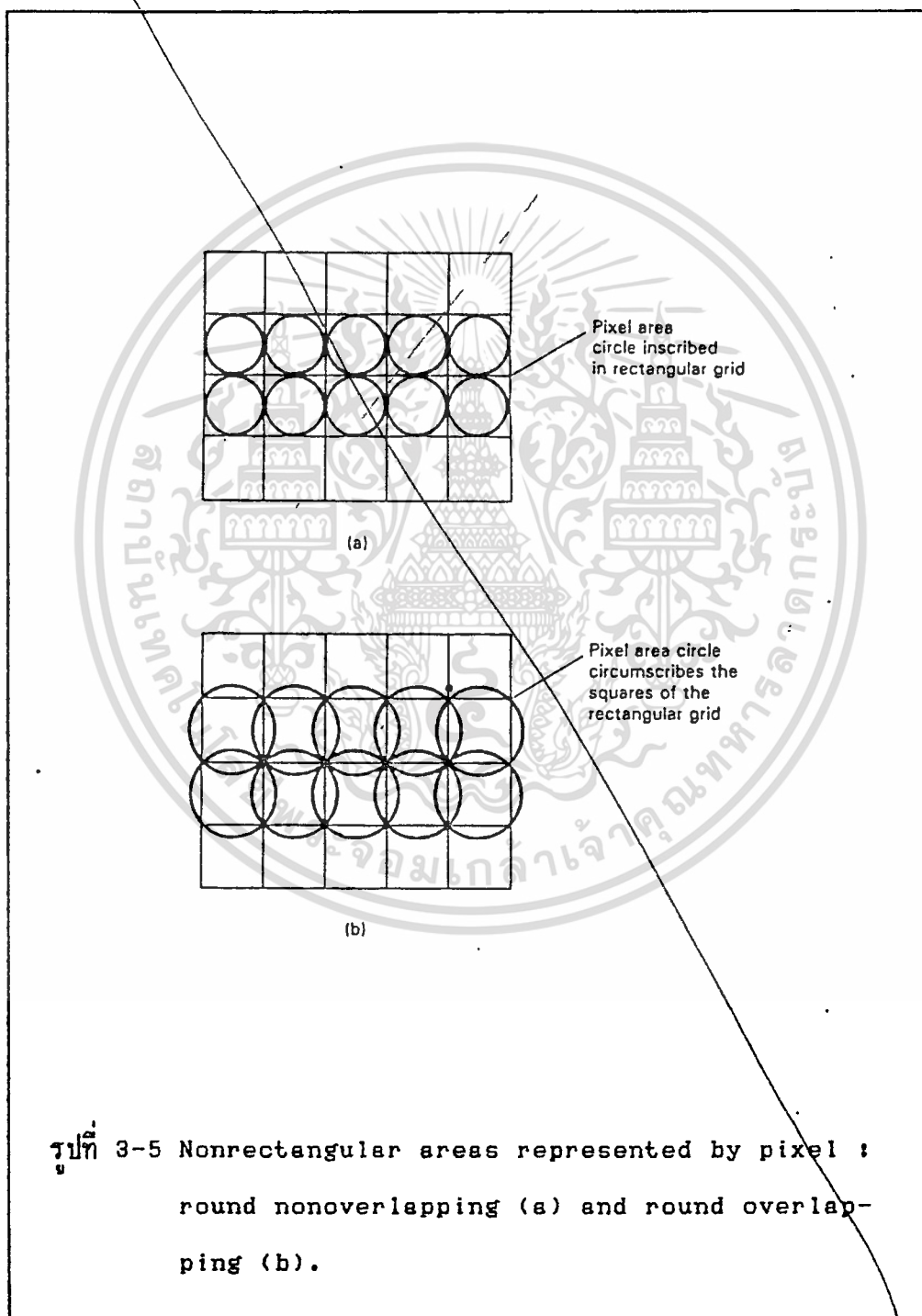
ในตัวอย่าง ส่วนของภาพ (picture element) เป็นรูปสี่เหลี่ยมผืนผ้า พื้นที่บนภาพจะเป็นรูปสี่เหลี่ยมผืนผ้าหรือวงกลมนั้นขึ้นอยู่กับตัวเซ็นเซอร์ ในกรณีที่เป็นกล้องที่ใช้หลอด (tube-camera) กับพื้นที่ตัวเซ็นเซอร์เป็นแนววงกลม อาจจะมีการเหลื่อมกัน (overlap) ดังแสดงใน

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูป 3-5

ในกรณีที่ไม่มี การเหลื่อมกัน (overlap) จะมีพื้นที่ของภาพระหว่างวงกลมซึ่งไม่ถูกวัดในกรณีที่มีการเหลื่อมกันของวงกลม จะมีพื้นที่ซึ่งถูกวัดเป็นสองครั้ง

ไม่มีข้อมูลใน pixel map ซึ่งแสดงถึงรูปร่างของพื้นที่ในภาพต้นแบบ (original picture) แทนโดยพิกเซล หรือข้อมูลเกี่ยวกับแสงสว่างใน picture element



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.4 Gray Scale

เพื่อที่จะให้มีข้อมูลที่อยู่ระหว่างค่าของแสงสว่าง จำเป็นจะต้องเพิ่มจำนวนบิตในการเก็บข้อมูลแทนค่าของพิกเซล ตัวอย่างเช่น ถ้าระดับของแสงสว่าง (illumination) อยู่ 4 ระดับจะต้องการจำนวนบิตในการเก็บข้อมูลเพียงสองบิต จำนวนบิตเพิ่มขึ้นเป็น 4 บิต จะมีระดับแสงสว่างได้ 16 ระดับ 8 บิต จะได้ 256 ระดับ จำนวนของระดับสีเทา (gray level) โดยปกติจะเป็นสองยกกำลัง n (2^n) ค่าที่อยู่ต่ำสุดคือ "0" ซึ่งจะถูกกำหนดให้เป็นสีดำ และ ตัวเลขที่เป็น "1" หรือมากกว่าขึ้นไปจะถูกแทนด้วยระดับสีเทาส่งสุด (maximum gray levels) ตัวอย่างเช่นค่า "15" ในระบบ 16 (ระบบที่มีระดับสีเทา 16 ระดับ) จะเป็นค่าสำหรับสีขาว การกำหนดค่าของพิกเซล ปกติมักจะเป็นเลขจำนวนเต็ม (integer)

GRAY SCALE	GRAY VALUE RANGE
2^1 2 ค่า	0, 1
2^2 4 ค่า	0 ถึง 7
2^4 16 ค่า	0 ถึง 16
2^8 256 ค่า	0 ถึง 256

การแทนความสว่างของแสงแบบไบนารีที่นิยมใช้กันอยู่ทุกวันนี้ในระบบการมองเห็น (vision system) เพราะมีข้อดีคือ ทำตัวเซ็นเซอร์ง่าย ข้อมูลมีความถูกต้องสูง การประมวลผลและการเก็บข้อมูล ซึ่งในปัจจุบัน ไมโครโปรเซสเซอร์มีการประมวลผลแบบใช้ข้อมูลมากกว่า 8 บิตแล้ว ดังนั้นระดับสีเทา (gray level) ที่มีค่าเท่ากับ 16, 64 และ 256 จึงเป็นเรื่องธรรมดา ในเวลานี้ได้จำกัดค่าไม่ให้เกิน 256 ระดับ เพราะว่า 64 และ 256 ระดับ ก็ให้รายละเอียดมากพอที่ตามมนุษย์จะสามารถเห็น และเพียงพอที่จะนำมาใช้ในงานอุตสาหกรรมได้อย่างพอเหมาะ ความสามารถของตามนุษย์ในการบอกความแตกต่างระหว่างความเข้มของแสงได้นั้น อยู่ในระดับ 40 แต่สามารถบอกความแตกต่างได้เด่นชัดเพียง 10-15 ระดับเท่านั้น ระดับสีเทามีค่าเท่ากับ 16 เป็นสเกลที่ต่ำสำหรับตามนุษย์ที่จะแยกแยะได้ ขณะที่ 64 และ 256 ระดับนั้น จะให้การแยกแยะได้มากกว่าตามนุษย์ที่จะแยกแยะได้

เอกสารระบบอาจจะมีความสามารถแบ่งได้ 256 ระดับ แต่ต้องประกอบพร้อมกับการประยุกต์ใช้งาน ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อาจจะทำให้จะต้องใช้รายละเอียดของจำนวนระดับสีเทา ความถูกต้องหรือความต้องการที่จะพัฒนาให้ดีขึ้น โดยใช้ $1/9, 1/13$ หรือ $1/20$ Subpixel resolution technique และระบบ Gray scale ของ 9, 13 หรือ 20 gray level scale ดังแสดงในรูป 3-6 จำนวนของ gray level มีผลต่ออิมเมจ โดยบริเวณที่มีแสงมากจะมีลักษณะเด่น (feature) ขณะที่แยกเฉพาะส่วนอธิบาย โดยทั่วไปการเพิ่มขึ้นของจำนวน gray level จะช่วยให้คุณภาพของอิมเมจดีขึ้นและจัดให้เหมาะสมที่จะทำให้ gray region ดีขึ้น ถ้าข้อมูล (information) เป็นฐานข้อมูล (data base) ขณะที่วิเคราะห์ระดับ gray scale image การจับ (acquisition) ภาพใน binary mode จะลดข้อมูลในการเก็บ แต่มีขีดจำกัดที่สามารถใช้เทคนิคการขยาย gray scale (gray scale expansion technique) ในการวิเคราะห์ประมวลผลของการกระทำ (operation) การเพิ่มขนาดของพิกเซลอิมเมจจากค่าเล็กๆ เช่น 32×32 ไปเป็น 250×250 จะเพิ่มความละเอียด และระดับรายละเอียดของอิมเมจ นี้เป็นข้อแตกต่างจากการ "zoom" ซึ่งเป็นการเพิ่มขนาดของพิกเซล

ฉบับที่ 3

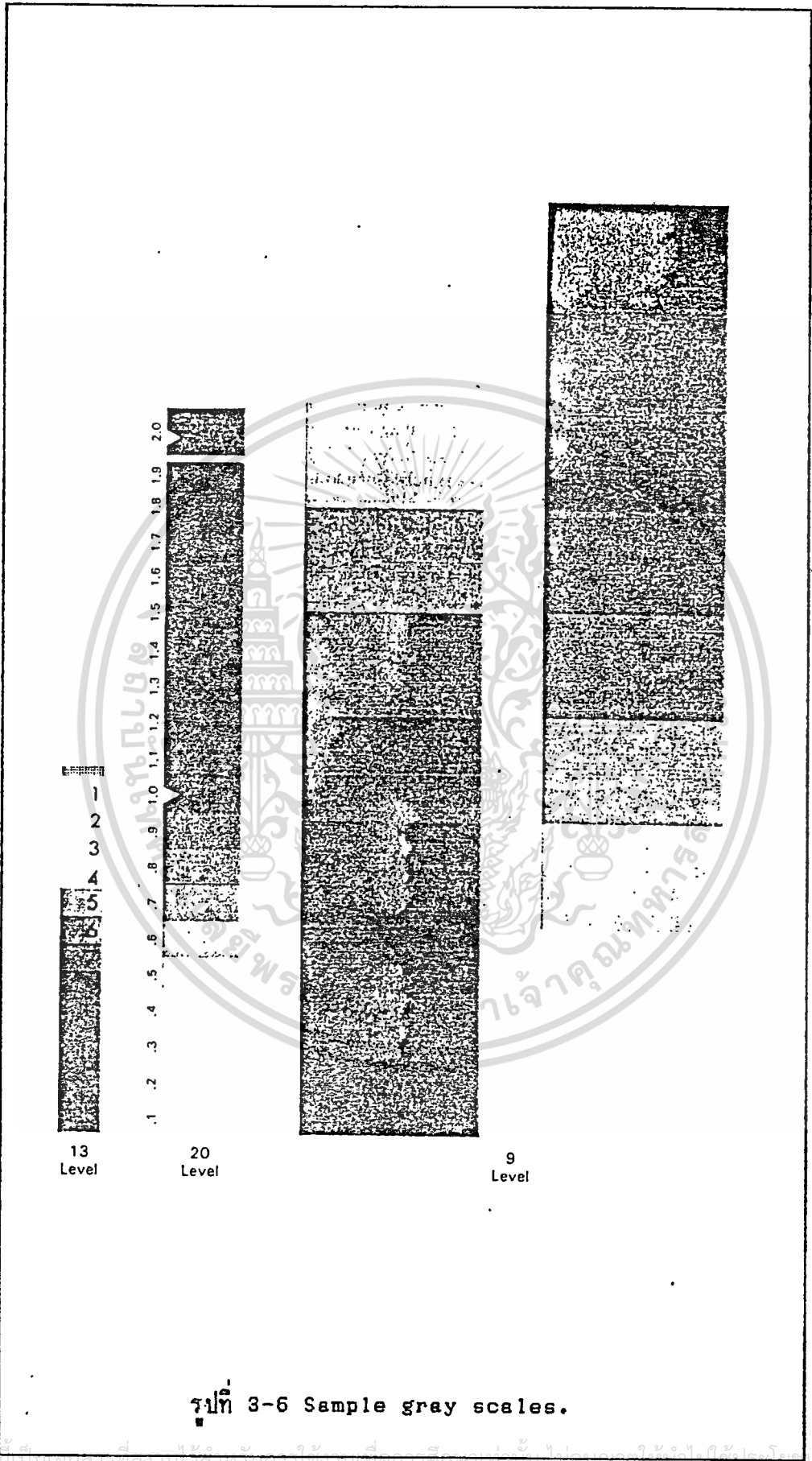
ในระบบ gray scale จะให้ค่ากลางๆ (means) ของลักษณะเฉพาะบอกความแตกต่างระดับของความเข้มแสงในรูปภาพ (picture) และให้ค่ากลางของเป้าหมาย subpixel ที่เที่ยงตรงในการวัดขนาดของนิพจน์ (item) back lighting จะถูกนำมาใช้ตรวจจับในส่วนของเครื่องจักรกับระบบการมองเห็น (vision system) สำหรับช่วงทั้งหมด (overall length) ในรูป 3-7

ส่วนของมุมใน diagram จะอยู่ในตำแหน่ง (2,1) , (2,4) , (5,1) , (5,5)

ในการวิเคราะห์แบบขังผึ้ง (cellular analysis) พื้นที่ที่แทนโดยพิกเซล (2,5) จะเป็นบริเวณมืด และ จะมีค่าเป็น "0" สำหรับที่ซึ่งไม่มีความเข้มของแสง พื้นที่ที่แทนโดยพิกเซล (5,2) จะมีวัตถุบังอยู่ครึ่งหนึ่ง ดังนั้น ค่าเฉลี่ยของแสงเหนือพื้นที่ดังกล่าวจึงเท่ากับ $(0+15)/2$ หรือ 7.5 พื้นที่ที่สว่างทั้งหมดจะมีค่าของพิกเซลเท่ากับ 15 ค่าความเข้มของแสงสว่างสูงสุดจะอยู่ในช่วง 0-15 ในระบบที่เป็น 16 ระดับ

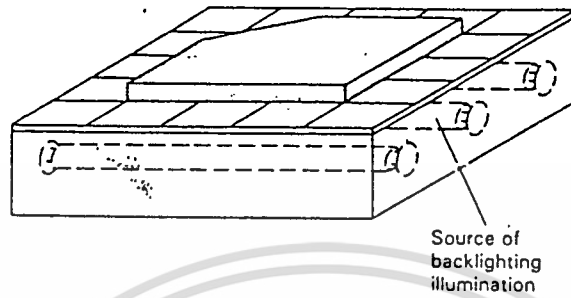
ขณะที่ค่าของพิกเซลจะต้องเป็นจำนวนเต็ม ค่า 7.5 จำเป็นต้อง modify การ modify จำเป็นจะต้องมีกฎเกณฑ์ ในระบบใดๆซึ่งจะหาว่า จะเปลี่ยน fractional value ไปเป็นจำนวนเต็ม (integer) ได้อย่างไร ตัวอย่างเช่น กฎเกณฑ์ทั่วไป จะปัดเศษขึ้นไปยังเลขจำนวนเต็มที่อยู่ถัดไป ถ้า fractional value มีค่าอยู่ระหว่าง 0.5 และ 1.0 ดังนั้น ถ้าวัดค่าได้เป็น 7.5 ค่าที่ใช้จะเป็น 8 ขณะที่ในรูป 3-7 ไม่ได้ปัดเศษขึ้น ถ้าวัดค่าได้ 6.6 ค่าที่ใช้ คือ 7 ผลของ fractional value ถูกปัดเศษขึ้นมีความสำคัญ เพราะมันมีผลกระทบต่อ tolerance และเป็นสิ่งสำคัญมากที่จะต้องรู้ว่า ค่าใดที่กำลังถูกใช้ในขณะนี้

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

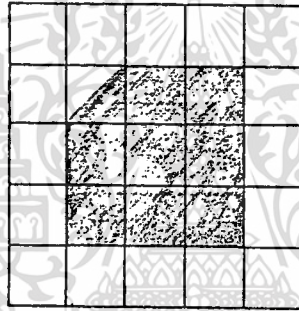


รูปที่ 3-6 Sample gray scales.

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อการใช้งานเท่านั้น ไม่อนุญาตให้ทำซ้ำโดยไม่ได้รับอนุญาตด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(a)



(b)

15	15	15	15	15
15	7	0	0	15
15	0	0	0	15
15	0	0	0	15
15	15	15	15	15

(c)

รูปที่ 3-7 Object on light table (a), with top view (b), and corresponding image matrix (c) where black pixel is represented by zero and white pixel is represented by fifteen.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับนักเรียนใช้เฉพาะเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.5 Quantizing Error

การตัดเศษค่า gray scale ไปเป็นจำนวนเต็ม จะได้ผลในความไม่แน่นอนของค่าที่ถูกต้อง ความไม่แน่นอนจะเป็นฟังก์ชันของกฎเกณฑ์การตัดเศษ

กฎเกณฑ์ซึ่งบอกว่า ค่าใดที่มีค่ามากกว่าค่าจำนวนเต็ม จะถูกปัดขึ้นเป็นจำนวนเต็มถัดไป ดังแสดงในรูป 3-8 สำหรับระบบที่มีค่า gray level เท่ากับ 16 การตัดเศษนี้ จะให้ผลในความไม่แน่นอนเชิงลบ (negative uncertainty)

ขนาด (dimension) และความกว้าง (tolerance) สามารถหาได้โดย การใช้ sub-pixel resolution technique และความรู้ในเรื่องการตัดเศษ ในตัวอย่างต่อไปนี้ ขนาด x (x dimension) ในรูป 3-9 ถูกแทนโดย 4 ช่องพิกเซล และ "1" ส่วนย่อยพิกเซลขนาด y (y dimension) ถูกแทนโดย 1 ช่องพิกเซล

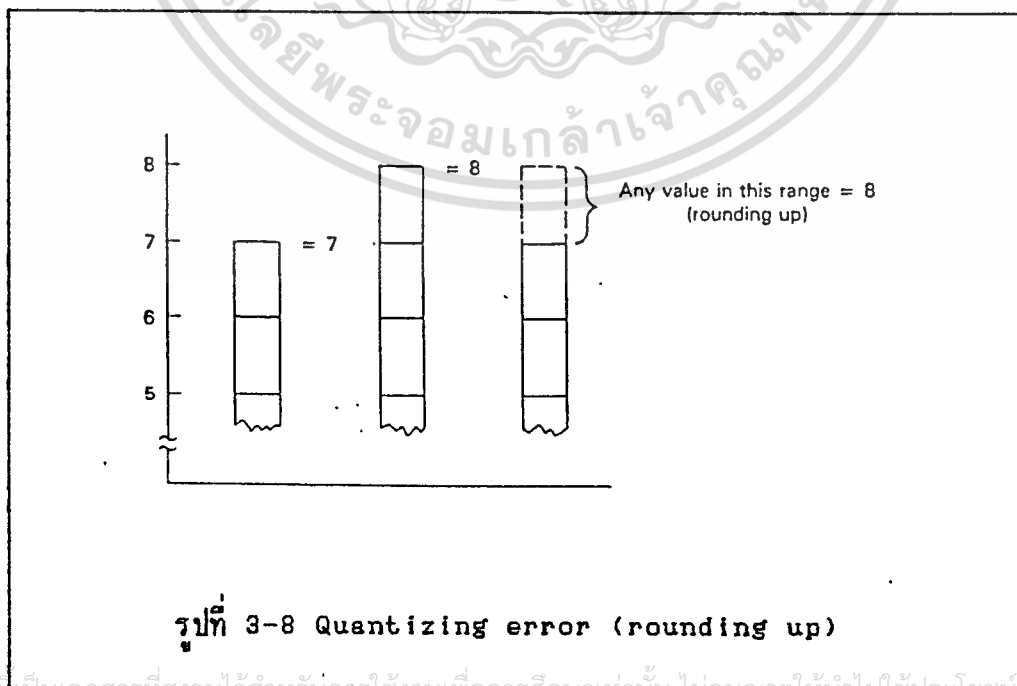
$$x = 4 \text{ pixel} \times 0.100 \text{ inch/pixel (on object)}$$

$$+ 1 \text{ partial pixel} \times \frac{12}{16} \times 0.100 \text{ inch/pixel}$$

12/16 factor คือ = proportion of gray levels in partial pixel/total gray levels.

$$x = 0.400 \text{ inch} + 0.075 \text{ inch} = 0.475 \text{ inch}$$

$$y = 1 \text{ pixel} \times 0.100 \text{ inch/pixel} = 0.100 \text{ inch}$$



เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้สำหรับใช้ภายในหน่วยงานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ทางการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

tolerance หาได้จากกฎการปิดเศษ ค่าพิกเซลของ "0" แทนด้วยค่าสูงสุดของแสงสว่าง ค่าพิกเซลของ 11 แทนด้วยค่าของแสงสว่างจาก 10.x ถึง 11.0 ซึ่ง x เป็นค่าจำนวนเต็มทีมากกว่าหรือ เท่ากับ "1" ค่าพิกเซลของ 15 ถูกแทนด้วยค่าแสงสว่างจาก 14.x ถึง 15.0 ซึ่งพอสรุปได้ดังนี้

Pixel value in image matrix		value of illumination represented by pixel value
0	=	0
1	=	$0 < \text{pixel value} \leq 1$
11	=	$0 < \text{pixel value} \leq 11$
15	=	$0 < \text{pixel value} \leq 15$

ดังนั้นการวัดที่คลาดเคลื่อนในพิกเซลกับการปิดเศษ ทำให้วัตถุไม่สามารถขยายได้มากกว่าที่แสดงให้เห็นโดยการคำนวณขนาด (dimension) แต่สามารถทำให้น้อยลงโดย การรวมของขนาดพิกเซล (pixel dimension) แทน โดย 1 gray level ในทิศทาง x จุดสิ้นสุดของสองพิกเซล จะช่วยให้การวัดไม่แน่นอน ขณะที่ในทิศทาง y มีเพียง 1 พิกเซลที่ทำให้เกิดความไม่แน่นอน (Uncertainty)

Tolerance or uncertainty of dimension

$$x = +0.0 \text{ inch} - 2 \text{ pixel} \times 1/16 \times 0.100 \text{ inch/pixel}$$

$$= -0.0125 \text{ inch}$$

$$y = +0.0 \text{ inch} - 1 \text{ pixel} \times 1/16 \times 0.100 \text{ inch/pixel}$$

$$= -0.00625 \text{ inch}$$

ขนาดของวัตถุ คือ

$$x = 0.475 + 0.0 \text{ inch} - 0.0125$$

เอกสารนี้เป็น 0.100 ที่ + 0.0 inch - 0.00625 ปรึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ค่าซึ่งเป็นค่าพิกเซลถูกเพิ่มขึ้นอีกหนึ่งของเลขจำนวนเต็มในการปัดเศษ และถูกลดค่าเมื่อปัดเศษลง (rounding down) ที่รู้จักกันในนาม "threshold" threshold จะอยู่ในช่วงระหว่างเลขจำนวนเต็มสองจำนวน จะเป็นผลทั้งในทางบวกและทางลบของ tolerance

กฎเกณฑ์ ซึ่งค่ามากกว่า 6.5 จะปัดเศษขึ้นเป็น 7 ค่าระหว่าง 0.6 และ 6.5 (จะปัดเศษลงเป็น 6.0 จะส่งผลทั้งทางบวกและทางลบของบริเวณไม่แน่นอน (uncertainty zone) ซึ่งเท่ากับ

1

$$2 \times (\text{max No. gray level})$$

จำนวน gray level ที่สูง จะทำให้เกิด quantizing error น้อย

3.5.1 Measurement Error

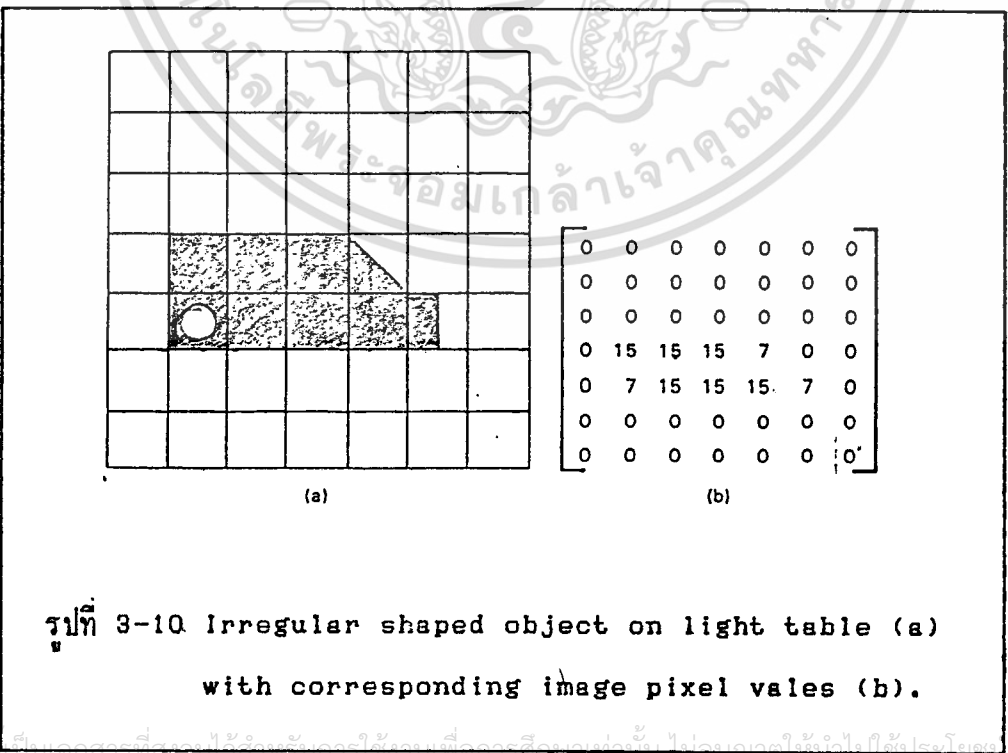
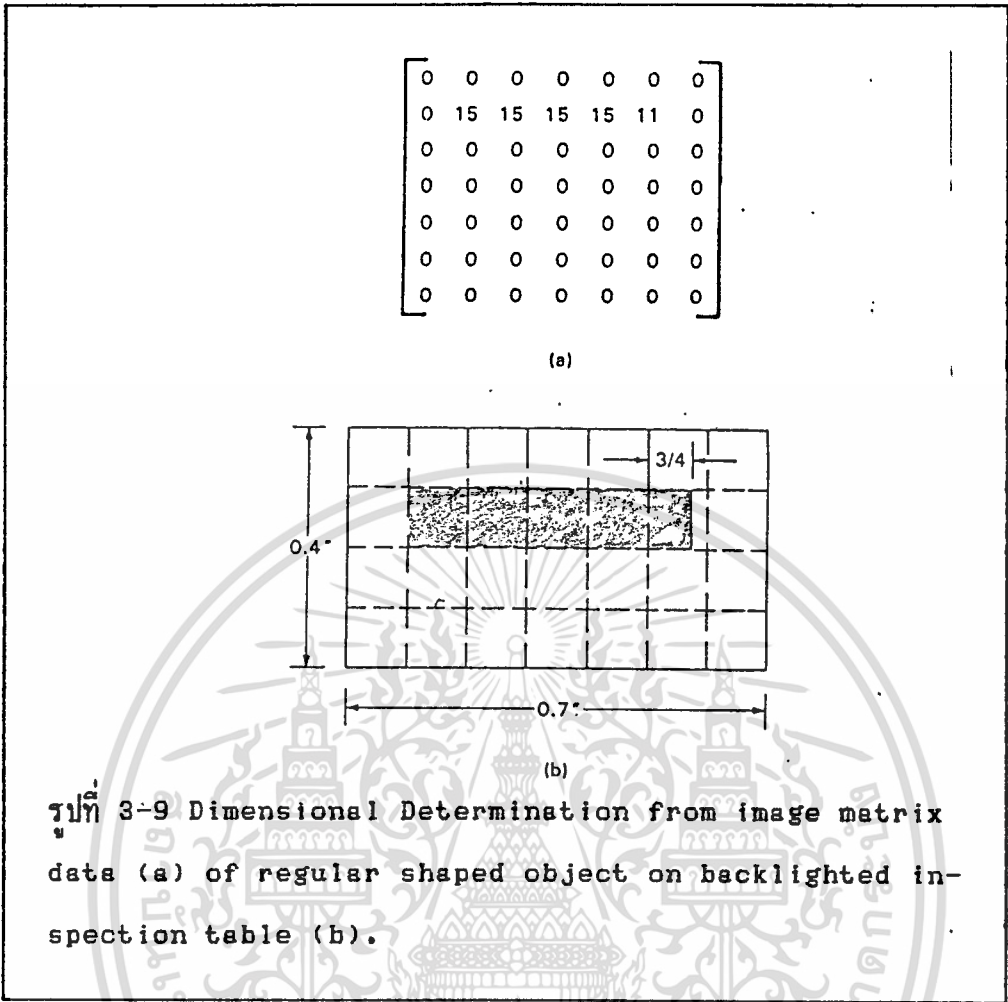
ขนาด (dimension) ของวัตถุในอิมเมจสามารถหาได้โดยการนับจำนวนพิกเซล หรือส่วนของพิกเซล (part of the pixel) ในทิศทางที่สนใจ ถ้า back lighting ถูกนำมาใช้ อิมเมจจะเป็นดำหรือขาวอย่างใดอย่างหนึ่ง และค่ากลางของค่า gray scale จะครอบคลุมตำแหน่งของพื้นที่พิกเซลโดยวัตถุ ความถูกต้องของการวัดขนาด (dimension) จะมีผลกระทบโดย quantizing error รวมอยู่ด้วยกับค่าพื้นที่ของพิกเซล ณ ตำแหน่งสิ้นสุดของวัตถุ

ขนาดของวงรีของวัตถุรูปทรงสี่เหลี่ยมผืนผ้าในอิมเมจ จาก back lighted ของระบบที่มีระดับสีเทา 16 ค่า (16 gray-level system) กับขนาดพิกเซลของ 0.10×0.10 นิ้ว หาได้โดยการบวกขนาดของพิกเซลที่มีค่า gray scale สูงสุด บวกกับส่วนย่อยของขนาดแทนด้วยค่ากลางๆของ gray scale ดังแสดงในรูป 3-9 ซึ่งใช้การปัดเศษขึ้น

มีข้อแตกต่างของวัตถุซึ่งจะเป็นผลในค่ากลางๆ ของ gray scale ในระบบ back lighting ดังแสดงในรูป 3-10

พื้นที่พิกเซล 3 ส่วน กับค่าของ 7 ที่แทนผลกระทบของรูปร่างที่แตกต่างกันทางคุณลักษณะ พื้นที่ (2,4) คือ 50% ปกคลุมโดยส่วนที่สิ้นสุดที่ทึบม 45 องศา พื้นที่ (3,5) คือ 50% ที่ถูกปกคลุมโดยส่วนที่ยื่นออกมาของสี่เหลี่ยมผืนผ้าของวัตถุ และพื้นที่ (1,3) คือ 50% ที่ปกคลุมโดยวัตถุที่เป็นรู (hole) สิ่งที่จะต้องจำใช้คือ ต้องรู้ลักษณะเด่นของวัตถุในการวิเคราะห์ ค่าอิมเมจ matrix สำหรับวิเคราะห์ dimension และ tolerance

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

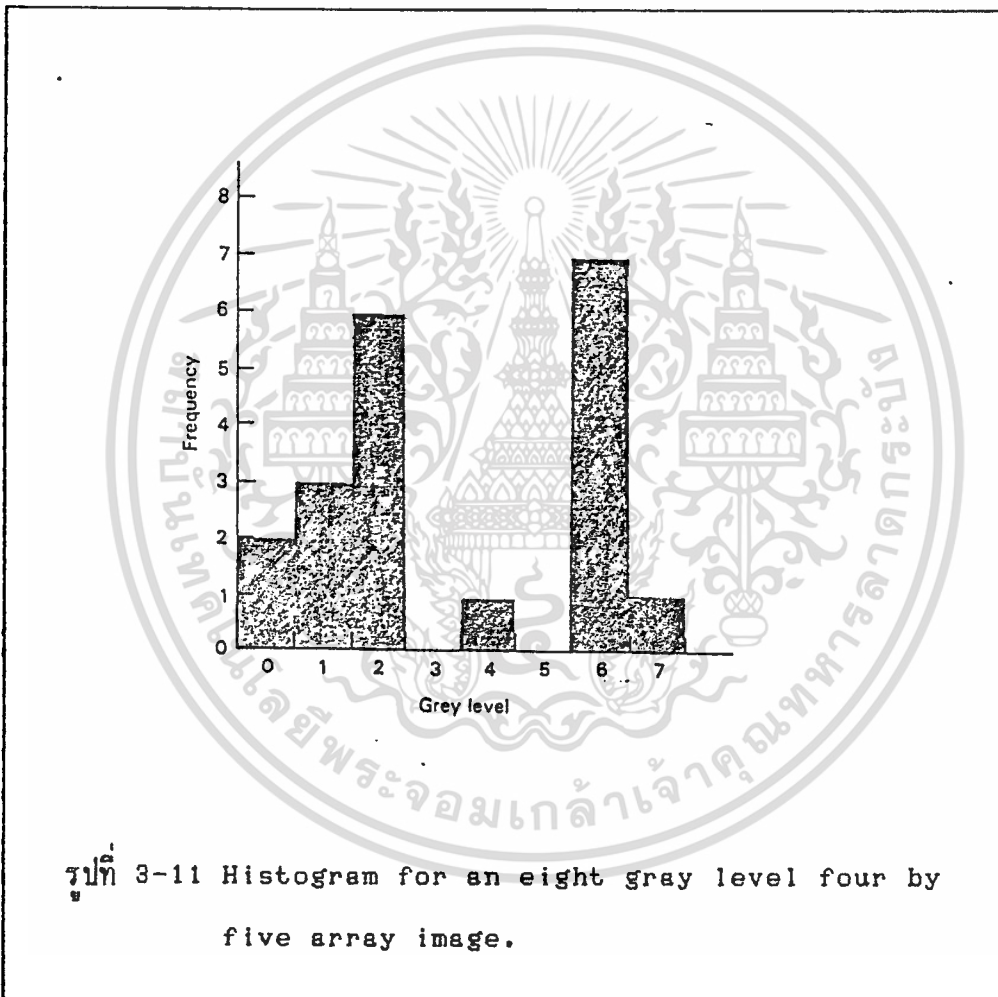


เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ซึ่งกรมที่ออกกรสิทธิ์นั้นจะไม่อนุญาตให้แก้ไขเปลี่ยนแปลงโดยไม่ได้รับอนุญาต
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.6 Histogram

Histogram คือ ภาพกราฟที่แทนความถี่ของการนับระดับสีเทา (gray level) ที่เกิดขึ้นในอิมเมจ แนวแกน x เป็นค่าของระดับสีเทา (gray level) และแนวแกน y เป็นจำนวนพิกเซลที่มีระดับสีเทา (รูป 3-11) Histogram ถูกสร้างโดย

- 1) Digitizing the image frame.
- 2) Counting the pixels at each gray scale level and
- 3) Plotting the frequency count of pixels at each gray level.



กราฟสามารถที่จะพิจารณารูปร่างแบบ bargraph ที่แสดงจำนวนของพิกเซลของแต่ละ gray level ที่เกิดขึ้นในกรอบ (frame)

ขนาด (magnitude) ของ histogram ที่ค่าของพิกเซล คือความน่าจะเป็นของค่าสีเทา (gray value) เกิดขึ้นในส่วนของภาพ (picture element) ในกรอบ (frame) ไม่มีข้อมูลที่ให้เกี่ยวกับตำแหน่งของพิกเซล ความน่าจะเป็นของพิกเซลค่า b เกิดขึ้น ณ ตำแหน่งจุด (x, y) ในรูปภาพ ดังรูป 3-11 และมิสมการใช้งานเพื่อการศึกษาเท่านั้น ไม่นุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

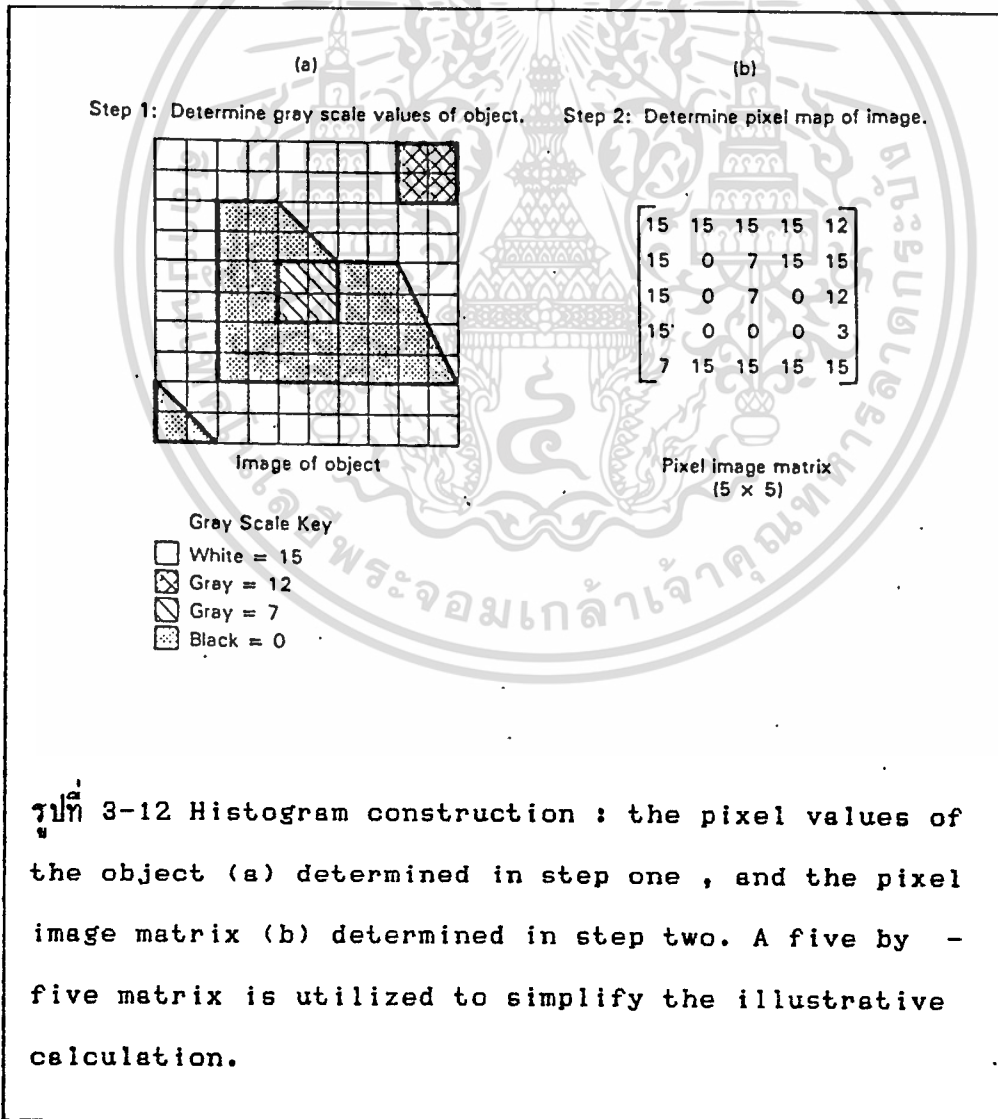
value b

$$P(b) = \text{at any point } (x,y) \text{ in the picture} = \frac{\text{total No. of pixels}}{\text{total No. of pixels}}$$

ถ้า $b = 6$ ระดับสีเทาและค่าของ histogram ที่ 6 เท่ากับ 7 คือ

$$P(6) = 7/20 = 0.35$$

รูปร่างของ histogram ที่เตรียมข้อมูลบน คุณลักษณะ (characteristic) ของภาพ ตัวอย่าง histogram ที่แคบบอกถึง contrast ในอิมเมจหรือรายละเอียดของค่าพิกเซลอาจแทนด้วย unique characteristic ของนิพจน์ (item) ในภาพ (picture) เช่นมีส่วนของรูป histogram มีประโยชน์ในการขีดค่า threshold ในการเปลี่ยน gray scale image ไปเป็นไบนารีอิมเมจ หรือ modify ส่วนของ gray scale spectrum



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.6.1 Construction

Histogram ของภาพสามารถสร้างขึ้นโดยขั้นตอนดังต่อไปนี้ ดังในรูป 3-12 และ 3-13

(1) หาจำนวนพิกเซลทั้งหมด ซึ่งจะใช้อธิบายภาพ นี้คือ การหาโดยอະเรย์ matrix ตัวอย่างนี้ใช้ matrix ขนาด $M \times N$ และจำนวนพิกเซลทั้งหมดเป็นผลจาก $M \times N$ ถ้า

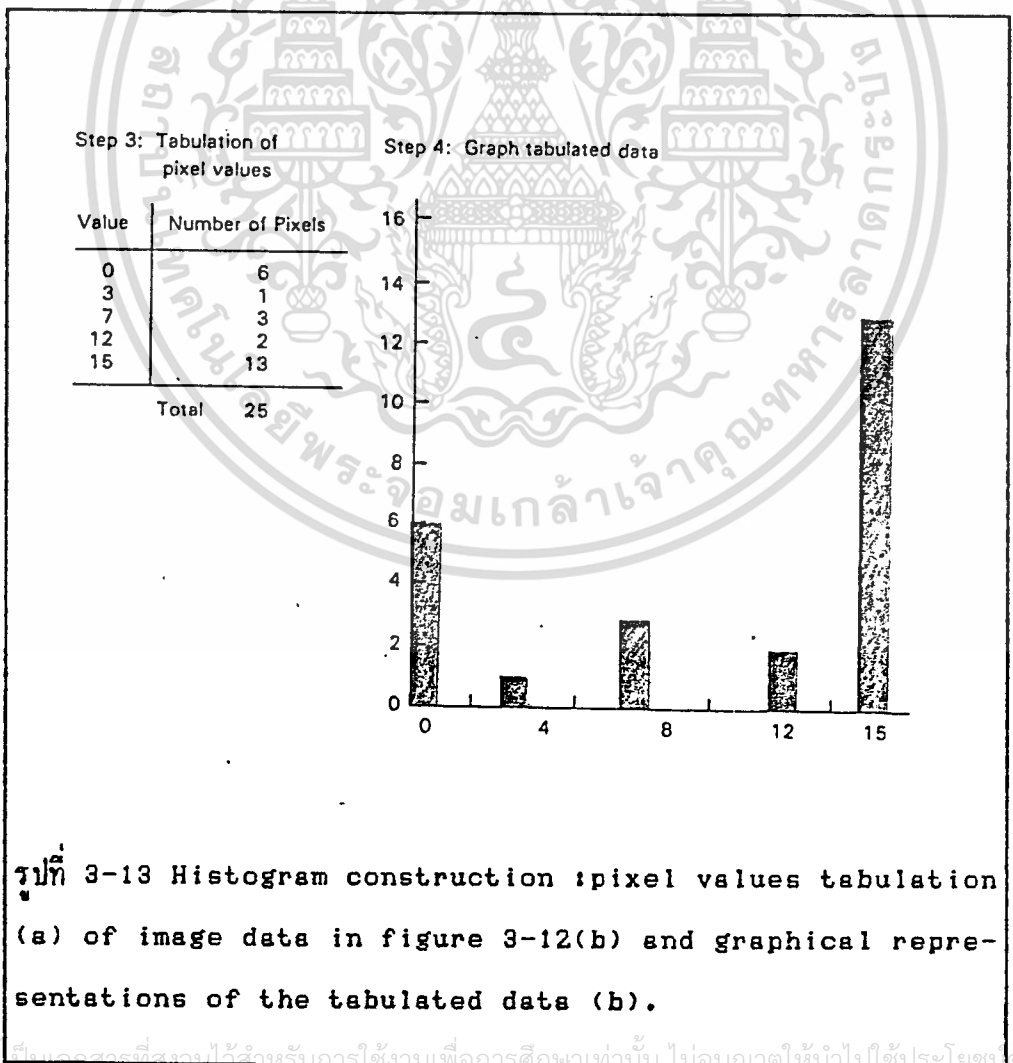
$$M = 10$$

$$N = 10$$

$$\text{Total pixel} = 10 \times 10 = 100$$

มันอาจเป็นไปได้ว่า M และ N ไม่เท่ากัน มันขึ้นอยู่กับชนิดของกล้อง และ analog-to-digital sampling rate และหน่วยความจำของระบบ ค่าที่สูงของ M และ N จะเพิ่มความเที่ยงตรง แต่ราคาก็จะเพิ่มขึ้นตามความจุและเวลาการตอบสนอง

(2) โครงสร้างของพิกเซลเม็บที่ (pixel map) ของอิมเมจในตัวอย่างเป็นโครงสร้าง matrix ขนาด 5×5 จำนวนพิกเซลทั้งหมด คือ 25



(3) ผลตารางของ pixel map นั่นคือ นับความถี่ของค่าพิกเซลแต่ละค่าที่เกิดขึ้น ตัวอย่างเช่น ค่า "0" เกิดขึ้นใน 6 พิกเซล แต่ค่า 3 เกิดขึ้นเพียงครั้งเดียว จำนวนของค่าทั้งหมดในตารางจะต้องเท่ากับ $N \times M$ ในตัวอย่างนี้เท่ากับ 25

(4) สร้าง histogram โดยพล็อต bargraph สำหรับค่าแต่ละพิกเซลจาก "0" ไปถึง gray scale ลบหนึ่ง ในตัวอย่าง gray scale ของ 16 ถูกใช้สร้าง pixel map และค่าสูงสุด คือ 15 มันจะแสดงค่าสูงสุดของพิกเซลใน map ไม่มีองค์ประกอบ (factor) histogram จะประกอบด้วย bargraph สำหรับค่าของพิกเซล ตั้งแต่ 0 ถึง 15



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

การประมวลผลการมองเห็นของเครื่องจักรเบื้องต้น
(Basic Machine Vision Processing)

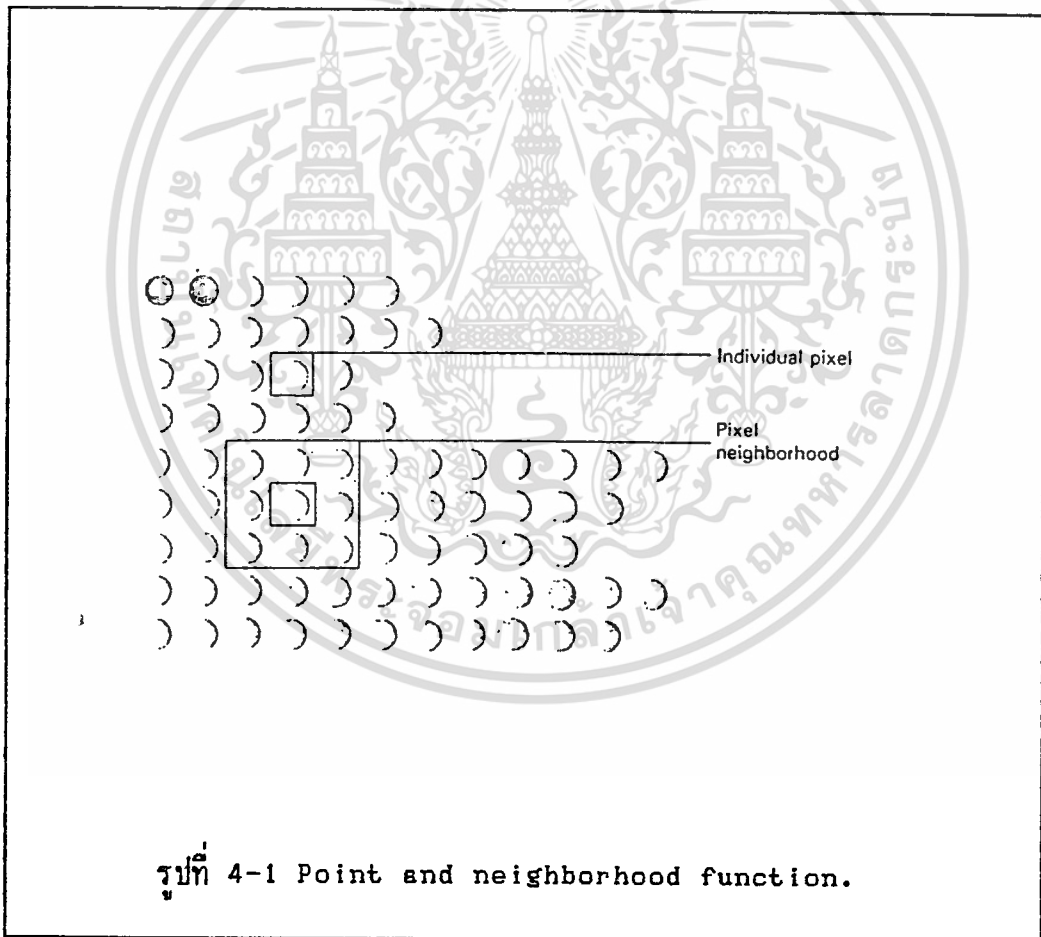
4.1 Introduction

การประมวลผลของข้อมูลในระบบการมองเห็นสามารถแบ่งได้ 2 ประเภท

(1) Point by point (monadic) การตัดแปลงข้อมูลบนสเกลทั้งหมด และ

(2) Multiple point (dyadic) การหา elements ของอะเรย์ตัวใหม่ของอิมเมจ

รูปแบบใหม่ของพิกเซลของอิมเมจเมทริก (image matrix) จะเป็นฟังก์ชันของค่าของตำแหน่งพิกเซลเดี่ยวๆ หรือค่าของพิกเซลในบริเวณข้างเคียงของเซลล์ (cell) เดี่ยวๆ ดังแสดงในรูป 4-1



Monadic operations เป็นการสร้างอะเรย์ตัวใหม่ โดยโมดิไฟ (modify) ค่าของพิกเซลที่ตำแหน่งบนพื้นฐานเดียวกันกับกฎเกณฑ์ทั้งหมด เอาไปใช้งานทุกๆตำแหน่งในอะเรย์ต้นแบบ การประมวลผลการหาค่าของพิกเซลในอะเรย์ ณ ตำแหน่งที่ให้ไว้ โมดิไฟ (modify) โดย simple linear operation หรือ nonlinear operation และใส่ค่าของพิกเซลใหม่ในตำแหน่งที่ใหม่

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

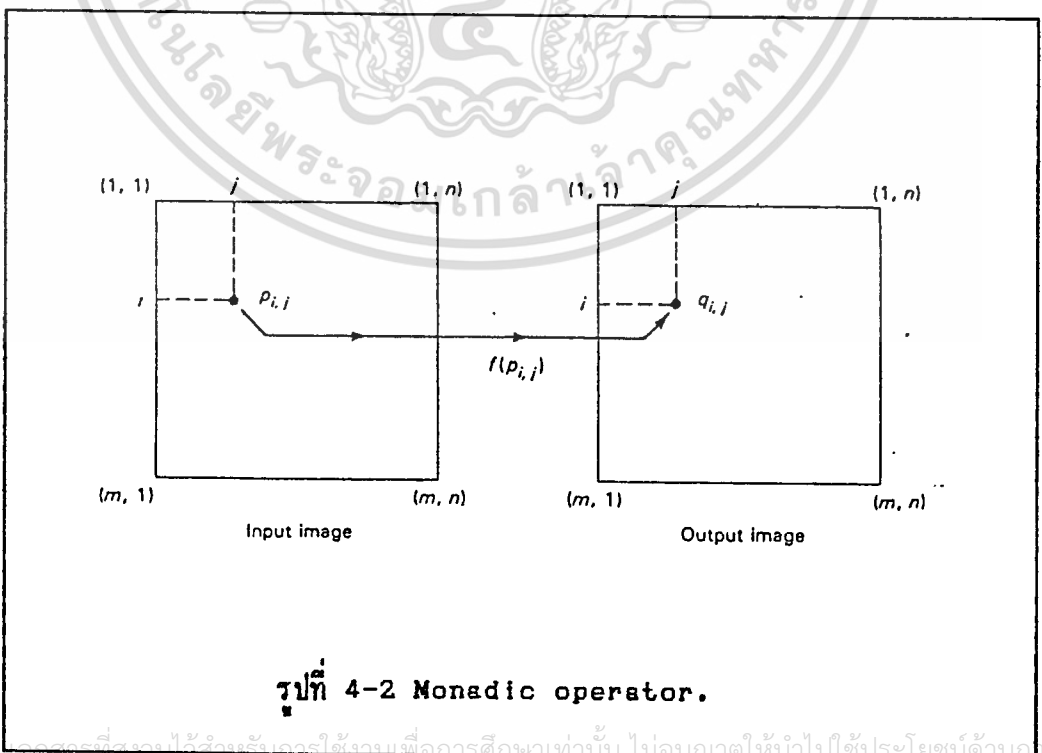
หน้าที่สอดคล้องกันในอะเรย์ตัวใหม่ การประมวลจะถูกทำซ้ำบนค่าของพิกเซลถัดไป และทำต่อไปจนครอบคลุมทั้งอะเรย์ (รูป 4-2) ที่แสดงให้เห็นในรูป 4-2 monadic operator เป็นแบบ one-to-one transformation ตัว operator f จะถูกใช้กับทุกพิกเซลในอิมเมจ หรือส่วนของอิมเมจและ output ขึ้นอยู่กับขนาด (magnitude) ของอินพุตพิกเซล (input pixel) ที่สอดคล้องกันเพียงตัวเดียว เอาท์พุท (output) ขึ้นอยู่กับพิกเซลข้างเคียง ฟังก์ชัน transform ค่า gray level ของทุกพิกเซลในกรอบ (frame) และค่าใหม่ถูกกำหนดโดยสมการ

$$g(i, j) = f[p(i, j)]$$

ฟังก์ชัน f สามารถเป็นได้ linear operator หรือ nonlinear operator คณิตศาสตร์ที่ใช้ในการประมวลผล เป็นความสัมพันธ์แบบธรรมดา ซึ่งเกี่ยวกับการหาค่าจากผลของสองค่า ผลของอิมเมจ matrix จะมีขนาดเท่ากับอิมเมจต้นแบบ

Dyadic operator ที่จำเป็นใช้ใน procedure ที่เหมือนกัน ซึ่งอะเรย์ของอิมเมจใหม่สามารถสร้างได้ 2 ทาง

- (1) ค่าของแต่ละพิกเซลในอะเรย์ขึ้นอยู่กับ การรวมค่าของพิกเซลที่สอดคล้องกันในตำแหน่งใน 2 หรือ มากกว่า image frame
- (2) ค่าพิกเซลในบริเวณที่อยู่ข้างเคียงที่ตำแหน่งของพิกเซล สร้างโดยการรวมกัน



Procedure หลังสุดจะทำให้สำเร็จได้โดยกรรมวิธี convolution การ convolution จะเกี่ยวข้องกับขนาดของ matrix แต่ 3×3 matrix เป็นขนาดที่ใช้กันโดยทั่วไป เพราะมันใช้เวลาในการคำนวณน้อย และคอมพิวเตอร์สามารถทำ matrix ขนาดใหญ่ๆได้ ในรูปแบบง่าย ๆ การ convolution ของสอง matrices , p และ q เกี่ยวกับการคูณของอีลิเมนต์ที่สอดคล้องกันทั้งหมด ผลของผลรวมของ element ที่ถูกคูณโดยค่าคงที่ค่าหนึ่ง (normalization - constant) เพื่อรักษาผลของค่าพิกเซลให้อยู่ภายในช่วงที่กำหนด ผลของค่าพิกเซลถูกลดแทรกในตำแหน่งที่เหมาะสมของอะเรย์ตัวใหม่ และ procedure ก็จะทำซ้ำหลังจากที่คู่ลำดับ (coordinate) ของแต่ละ element จะถูกเลื่อน (shift) ไปอีกหนึ่งในทิศทาง M หรือ N อย่างใดอย่างหนึ่ง จนกระทั่งตำแหน่งทั้งหมดในอะเรย์ถูกทำจนหมด

4.2 Monadic one point transformation

Monadic point-by-point operators เป็นวิธีที่ง่ายที่สุดในการประมวลผลภาพ (image processing) กรรมวิธีจะเกี่ยวกับ input image เพียง 1 อิมเมจ (p) ดังแสดงในรูป 4-2 และผลที่ได้ก็เป็น single image output , Point-to-point operator ซึ่งจะกล่าวถึงในส่วนนี้จะ เป็น identity operator , inverse operator , threshold operator และ combination

4.2.1 Identity Operator

Operator ตัวนี้จะให้ผลในการสร้างเอาท์พุทอิมเมจ ซึ่งเป็นอิมเมจแบบเอกลักษณ์ (identical) กับอินพุทอิมเมจ ดังแสดงในรูป 4-3 ค่าของแต่ละพิกเซลในอิมเมจที่สอง (q) เป็นเอกลักษณ์ (identical) กับค่าของพิกเซล p ที่สอดคล้องกันในอิมเมจที่หนึ่ง ฟังก์ชัน f เป็นเส้นตรงที่มีจุดเริ่มต้นที่จุดกำเนิดและต่อออกไปถึงค่าพิกเซลสูงสุดของระบบอิมเมจ procedure พื้นฐานในการประมวลผลภาพ และ operator ตัวอื่นๆ จะทำให้สำเร็จได้โดยการ modify ฟังก์ชัน

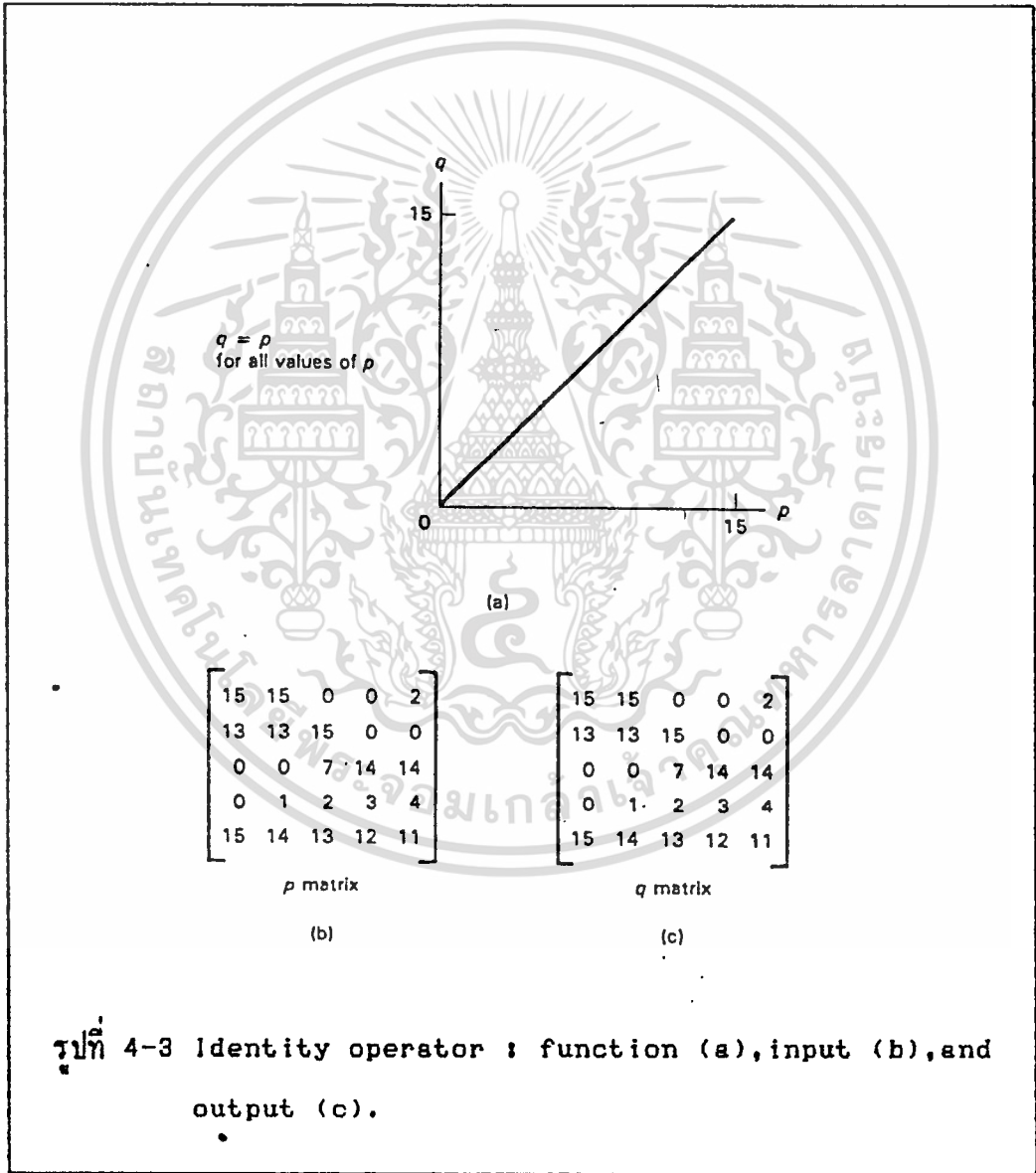
4.2.2 Inverse Operator

Operator ตัวนี้จะให้ผลในการสร้างเอาท์พุทอิมเมจซึ่งเป็นส่วนกลับ (inverse) กับอินพุทอิมเมจ วิธีการก็คล้ายกับ identity operator ยกเว้นค่าฟังก์ชันที่แตกต่างกัน ดังแสดงในรูป 4-4 ฟังก์ชัน f เป็นเส้นตรงที่มีค่า gray level สูงสุดที่ค่า gray input ต่ำสุด และเท่ากับ "0" ที่ค่า gray input สูงสุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.3 Threshold Operator

Operator ในขั้นนี้จะให้ผลเป็นไบนารีเอาท์พุทอิมเมจ จาก gray scale output image ซึ่งระดับของการส่งผ่านที่ให้โดยอินพุทพารามิเตอร์ p , ดังแสดงในรูป 4-5 และรู้จักกันในนาม "threshold" ค่าของพิกเซลทั้งหมดที่ต่ำกว่า p , ถูกเปลี่ยนเป็น "0" และค่าของพิกเซลทั้งหมดที่เท่ากับหรือมากกว่า p , จะถูกเปลี่ยนเป็น "1" operator ตัวนี้สามารถใช้ในการหา spatial information จาก histogram โดยทำ procedure ซ้ำ โดยใช้ค่า threshold ที่แตกต่างกัน



4.2.4 Other Operator

Operator ที่แตกต่างกันสามารถพัฒนา โดยการรวม operator พื้นฐาน 3 ตัว คือ identity, inverse และ threshold การศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Inverse threshold operator

ไบนารีเอาน้ำพุทิมเมจสามารถที่จะกลับได้โดยประยุกต์ฟังก์ชัน threshold ในรูป 4-6 ไปเป็นอิมเมจค่า threshold ในรูป 4-6 แสดงให้เห็นโดยค่า P_1 และตัวอย่างการใช้ค่า threshold ของ 5 จุดที่แสดงซึ่งเป็นค่าที่ถูกนับรวมทั้งหมดทุกพิกเซลในภาพต้นแบบ ซึ่งที่โหนดสว่างจะกลายเป็นมืด และที่โหนดมืดจะกลายเป็นสว่าง operator ตัวนี้มีค่าได้มากที่สุด เพราะค่า threshold สามารถที่จะแปรเปลี่ยนไปเพื่อหา spatial information

Binary Threshold Interval

operator ในขั้นนี้จะให้ผลเป็นไบนารีเอาน้ำพุทิมเมจ ซึ่งค่า gray input ในช่วง P_1 ถึง P_2 รวมทั้งค่า P_1 และ P_2 ด้วย จะถูกเปลี่ยนเป็น "1" และค่าที่อยู่นอกเหนือจากช่วง P_1 ถึง P_2 จะถูกเปลี่ยนให้เป็น "0" ดังแสดงให้เห็นในรูป 4-7 operator ตัวนี้สามารถนำไปใช้หา spatial information จาก histogram เพราะมันแสดงตำแหน่งของพิกเซลที่อยู่ในช่วงที่กำหนด

Inverse Binary Threshold Operator

operator ตัวนี้สามารถนำไปใช้เปลี่ยนอิมเมจที่มีค่า gray level หลายๆระดับ (multi level gray scale image) ไปเป็นไบนารีอิมเมจได้ หรือ มันสามารถประยุกต์ใช้กับไบนารีอิมเมจสำหรับวัตถุประสงค์ในทางกลับกัน

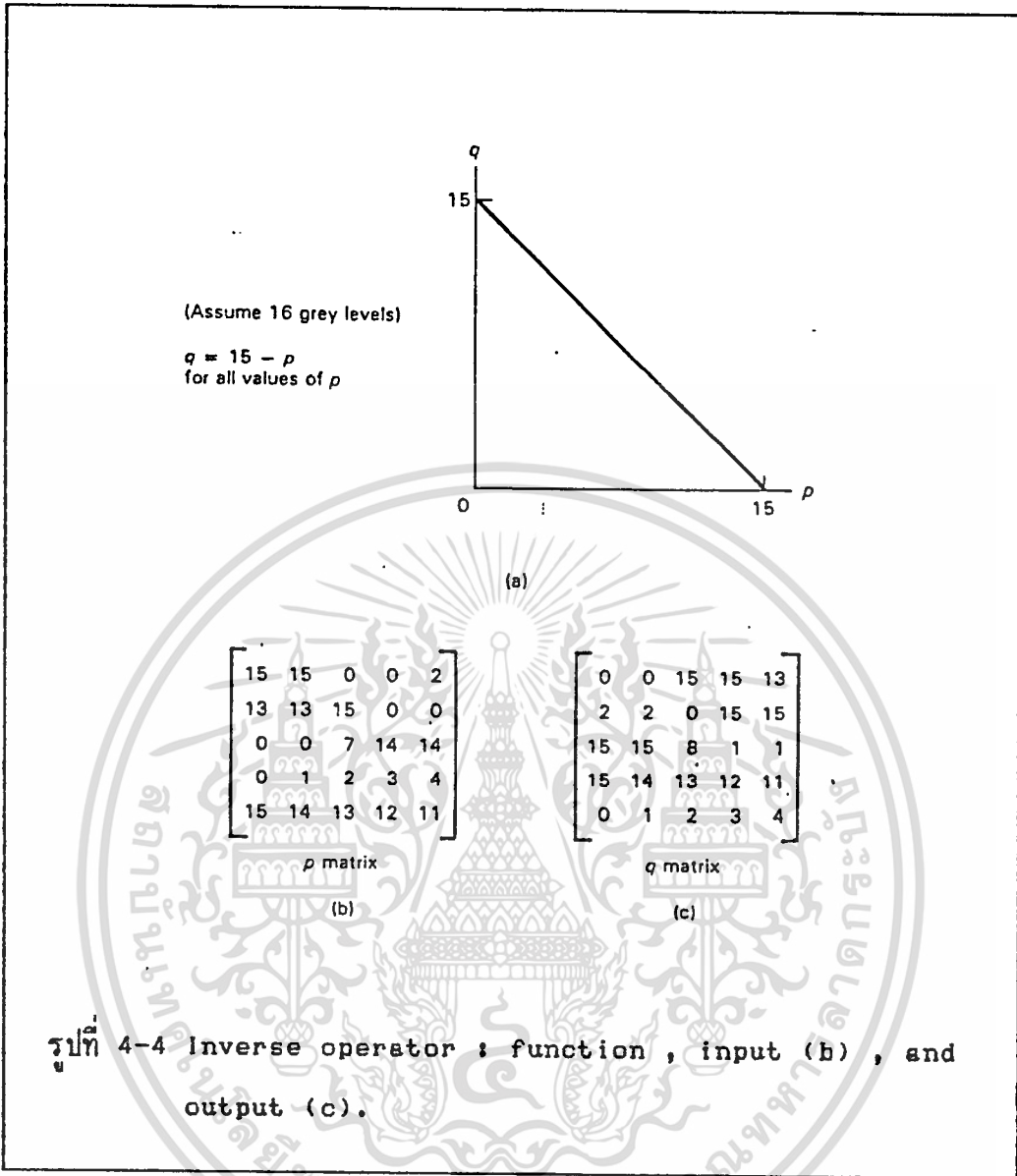
ไบนารีเอาน้ำพุทิมเมจ หรืออิมเมจที่ได้จาก Binary threshold interval operator สามารถที่จะกลับโดยกลับฟังก์ชัน Binary threshold ดังแสดงในรูป 4-8

Gray scale threshold operator

operator ในขั้นนี้แสดงไว้ในรูป 4-9 ผลใน gray scale output image สำหรับค่า gray scale image ระหว่าง P_1 และ P_2 และทำค่าที่อยู่นอกเหนือจาก P_1 ถึง P_2 ให้เป็น "0"

Inverse gray scale threshold operator

เอาน้ำพุทิมเมจสามารถที่จะกลับได้โดยใช้ฟังก์ชันที่พิสูจน์ให้เห็นโดยรูป 4-10 operator ตัวนี้ สามารถที่จะให้แสงสว่างในรายละเอียดที่มีลักษณะเด่น ตัวอย่างเช่น ถนน หรือ พื้นที่มีเอเหมือนกันในภาพที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

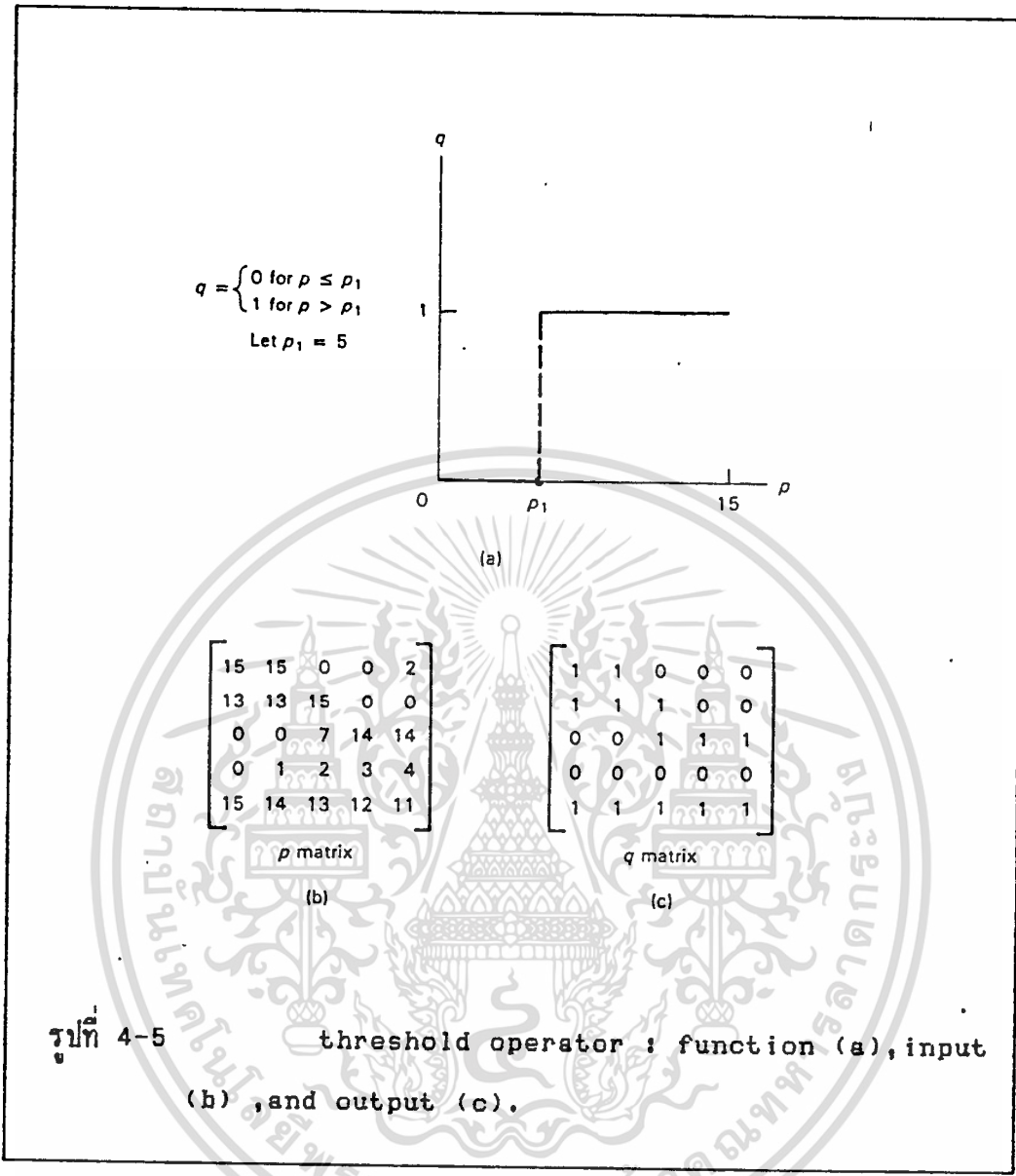


Stretch Operator

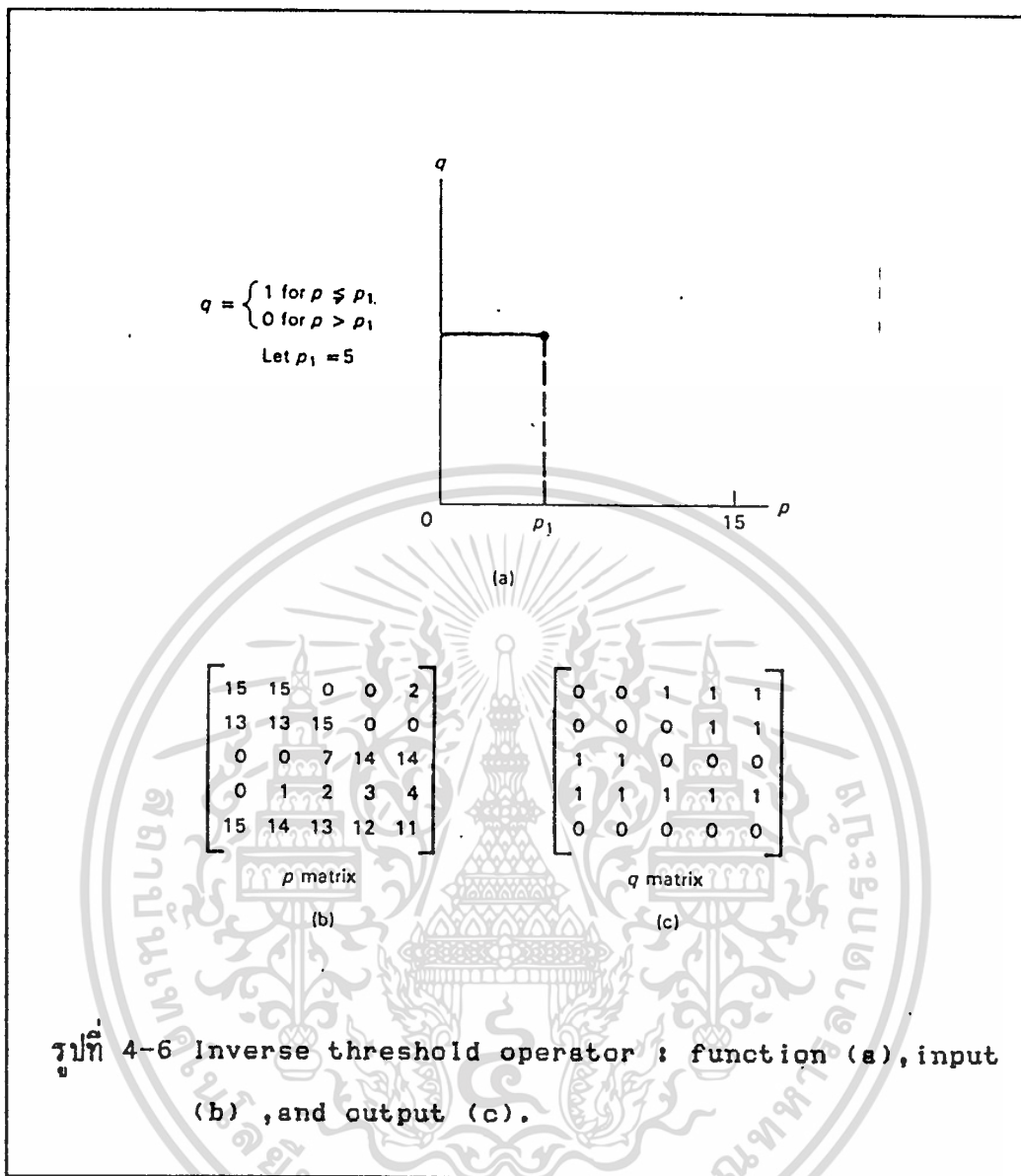
operator ตัวนี้จะให้ผลใน full gray scale output image ที่สอดคล้องกับช่วงอินพุต P_1 และ P_2 และกด (suppress) ค่าทั้งหมดที่อยู่นอกเหนือจากช่วงนี้ แสดงให้เห็นในรูป 4-11

Gray scale reduction operator

operator ตัวนี้ให้ผลในเอาท์พุทอิมเมจ ซึ่งมีจำนวนค่า gray level น้อยกว่าจำนวนของ gray level ของอินพุทอิมเมจ ดังแสดงในรูป 4-12 ซึ่งอินพุทอิมเมจกับค่าระดับสีเทา 15 ระดับจะถูกเปลี่ยนเป็นอิมเมจที่มีระดับสีเทา 5 ระดับ คือ q_1, q_2 ให้ q_3 ใช้ q_4 และ q_5 15รค่าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



4.3 Dyadic Two Point Transformation

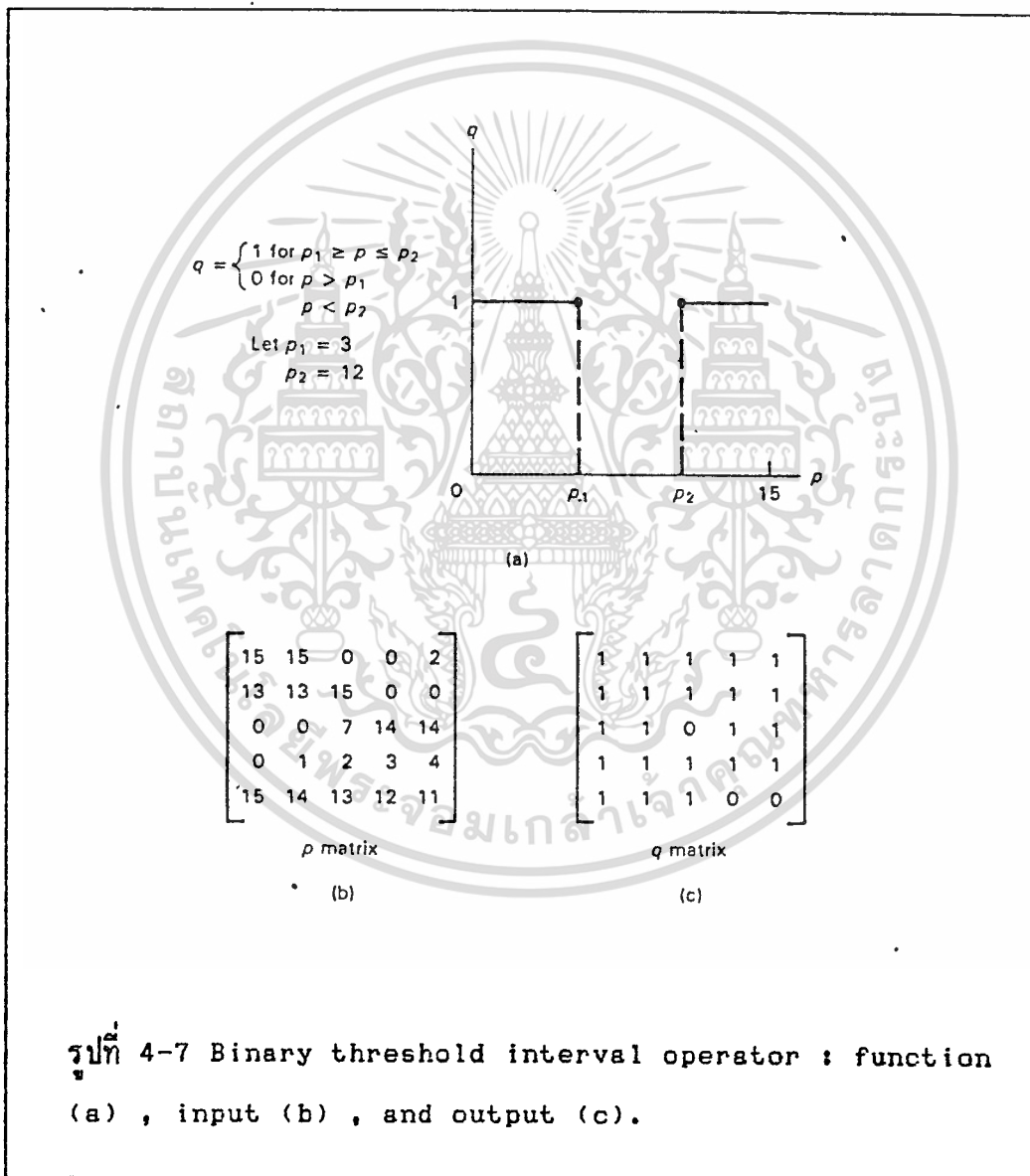
Dyadic point-by-point operator ใช้กับข้อมูลที่มีตำแหน่งที่เหมือนกันในสองอิมเมจ two-input image matrix ในรูป 4-13 พิกเซลจาก A และ B ถูกใช้สร้างอิมเมจใหม่คือ C ขนาดของ matrix ไม่เปลี่ยนแปลง และฟังก์ชัน dyadic transformation f_D สามารถเป็นได้ทั้งแบบ linear หรือ nonlinear

ฟังก์ชัน transformation ถูกนำไปใช้งานกับคู่ตำแหน่งของพิกเซลทั้งหมดอินพุทอิมเมจ (อินพุทอิมเมจมี 2 อิมเมจ) นั่นคือ ข้อมูลจากตำแหน่งพิกเซลในอิมเมจที่ 1 จะถูกรวมกับข้อมูลที่ตำแหน่งพิกเซลที่สอดคล้องกันในอิมเมจที่ 2 เพื่อสร้างค่าของพิกเซลที่สอดคล้องกันในอิมเมจที่ 3

เอกสารนี้ถูกจัดพิมพ์ขึ้นเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$C_{i,j} = f_D (a_{i,j}, b_{i,j})$$

ซึ่ง f_D เป็นฟังก์ชันของสองตัวแปร และ i กับ j เป็นช่วงจาก 0 ถึง m และ 0 ถึง n f_D สามารถที่จะบวก ลบ คูณหาร ยกกำลังได้สูงสุด หรือ ฟังก์ชันอื่นซึ่งสามารถถูกหารได้ แต่ต้องระวังว่า อิมเมจทั้งสองนั้น จะต้องอยู่กับที่และคุณสมบัติการ calibration และ registration ของสองอินพุทอิมเมจ เพราะการคำนวณจะอยู่บนพื้นฐานความสอดคล้องของจุดต่อจุดในการหา



รูปที่ 4-7 Binary threshold interval operator : function (a) , input (b) , and output (c).

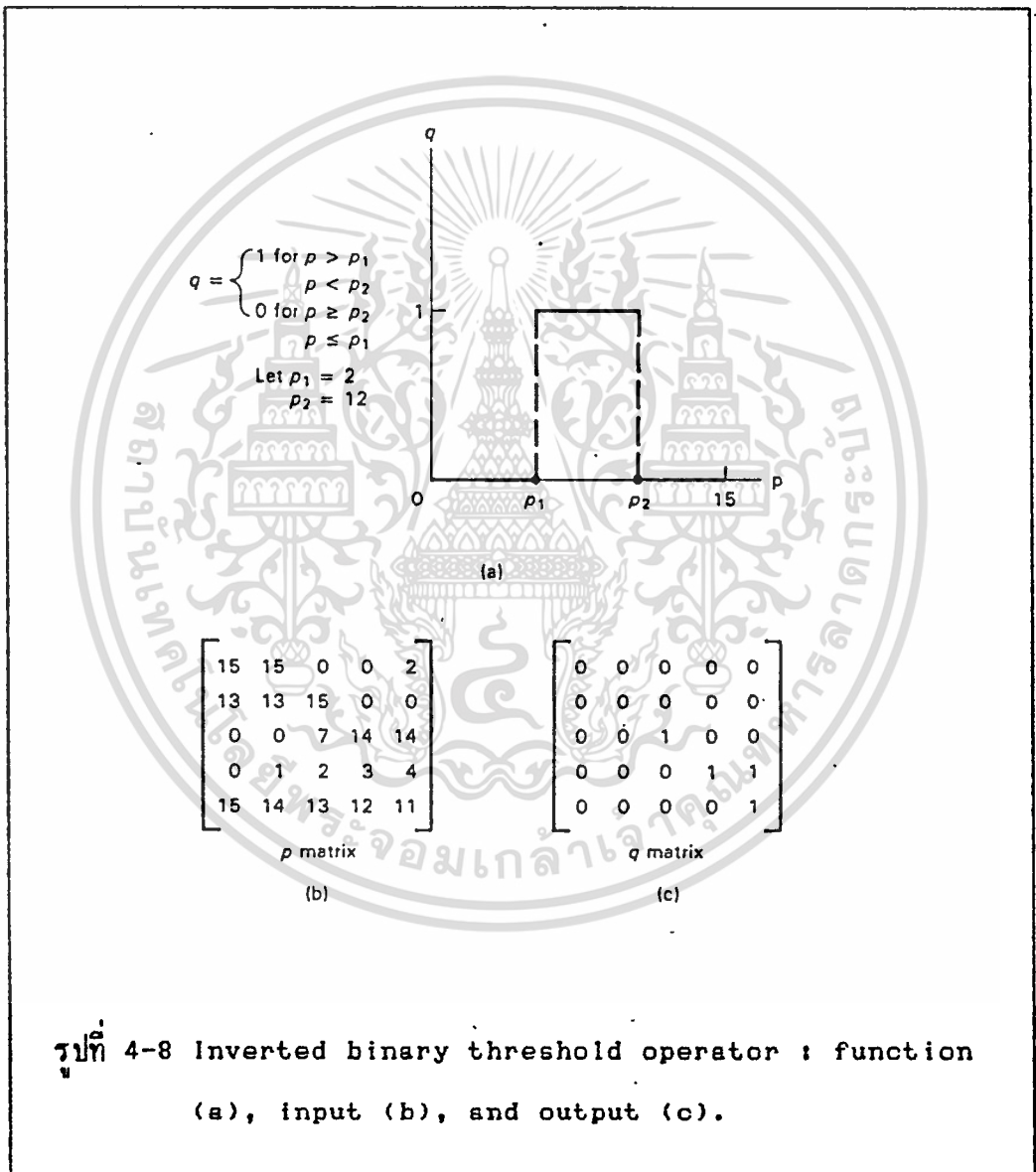
ฟังก์ชันจะมีค่า factor K เป็นค่า scaling ที่เหมาะสม เพื่อรักษาขนาด (magnitude) ของค่าเอาต์พุทให้อยู่ภายในช่วง ไม่ให้เกิด overflow หรือสภาวะเป็นลบ (negative condition)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Transformation เกี่ยวข้องกับสองตัวแปร รวมทั้งคู่ของพิกเซลที่สอดคล้องกัน

$$R(i, j) = f[P(i, j) , q(i, j)]$$

เมื่อ p และ q เป็น input matrix f เป็น functional operator และ R เป็นผลของ output matrix



4.3.1 Image Addition

Image addition สามารถใช้ลดผลกระทบของสัญญาณรบกวน (noise) ในข้อมูลแสดงไว้ในรูป 4-14 ค่าของเอาต์พุต $c_{i,j}$ ซึ่งให้โดยสมการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษานั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$c_{i,j} = (a_{i,j} + b_{i,j}) / K$$

i, j เป็นค่าของช่วง ซึ่งค่า K เท่ากับจำนวนลุ่ม (samples) image addition dyadic จะทำการประมวลผลค่าเฉลี่ยข้อมูลในสอง input image matrices ถ้าหนึ่งในอินพุตอิมเมจเป็นค่าคงที่ ผลที่ได้จะสว่างขึ้นหรือมืดลงทั้งอิมเมจ ซึ่งจะปรากฏขณะเลื่อน (shift) histogram

ในตัวอย่าง special rule เหมือนการปิดเศษขึ้น จะต้องถูกหารและนำไปใช้กับฟังก์ชัน transformation ที่ให้ผลค่ากลางๆ มีความเที่ยงตรงที่จะทำให้ดีขึ้น และลดสัญญาณรบกวน (noise) เมื่อ procedure ถูกใช้กับจำนวน samples ที่มาก

$$C_{i,j} = (a_{i,j} + b_{i,j}) / 2$$

i, j เป็นช่วงของค่าจะถูกปิดขึ้น

4.3.2 Image Subtraction

การลบอิมเมจสามารถใช้ในการตรวจจับ (detect) การเฉลี่ยที่เกิดขึ้นระหว่างช่วงเวลา ระหว่างสองอิมเมจที่ถูกถ่ายเข้าคนละเวลา ถ้าทั้งสองอิมเมจเป็นฉาก (scene) เหมือนข้อมูลอาจจะแทนด้วย heat loses หรือ cooling ถ้า infrared spectrum เป็นข้อมูลต้น เพราะว่าการประมวลผลภาพ (image processing) ใช้เลขจำนวนบวก มันจำเป็นต้องกำหนดเอาที่พิก ซึ่งทำให้เป็นค่าบวกทั้งหมด นี้ยังสามารถปรับสเกลได้ ซึ่งค่าที่ติดลบก็จะ set ให้เป็น "0" ค่าที่ใหญ่มากก็ set ให้เท่ากับค่าสูงสุดของ gray scale level (255 ในส่วนของรูป 4-15) อีกสิ่งหนึ่งที่ต้องทำให้เหมาะสม คือ การกำหนดความแตกต่างแบบไม่คิดเครื่องหมาย หรือ ค่า absolute ของความแตกต่างดังแสดงใน matrix b

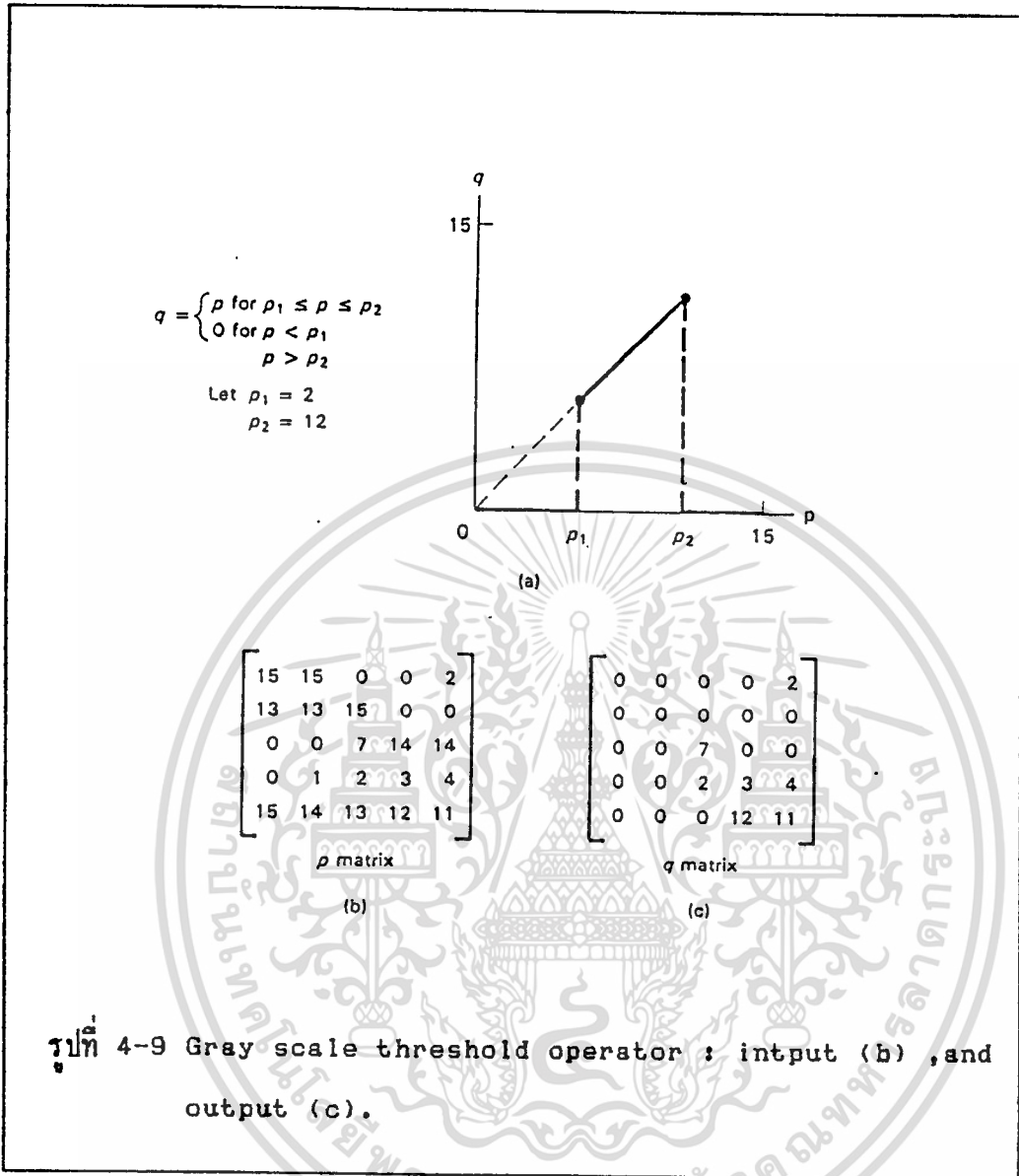
ในกรณีแรก (ตัวอย่าง a) มีความสัมพันธ์เป็น

$$C_{i,j} = K (a_{i,j} + b_{i,j})$$

เมื่อ K เป็นฟังก์ชันแบบไม่เชิงเส้น ซึ่งค่าต่ำสุดของ $C_{i,j}$ เป็น "0" และค่าสูงสุดคือ 255

ในตัวอย่าง b ของรูป 4-15 มีความสัมพันธ์เป็น

$$R_{i,j} = K (a_{i,j} - b_{i,j})$$

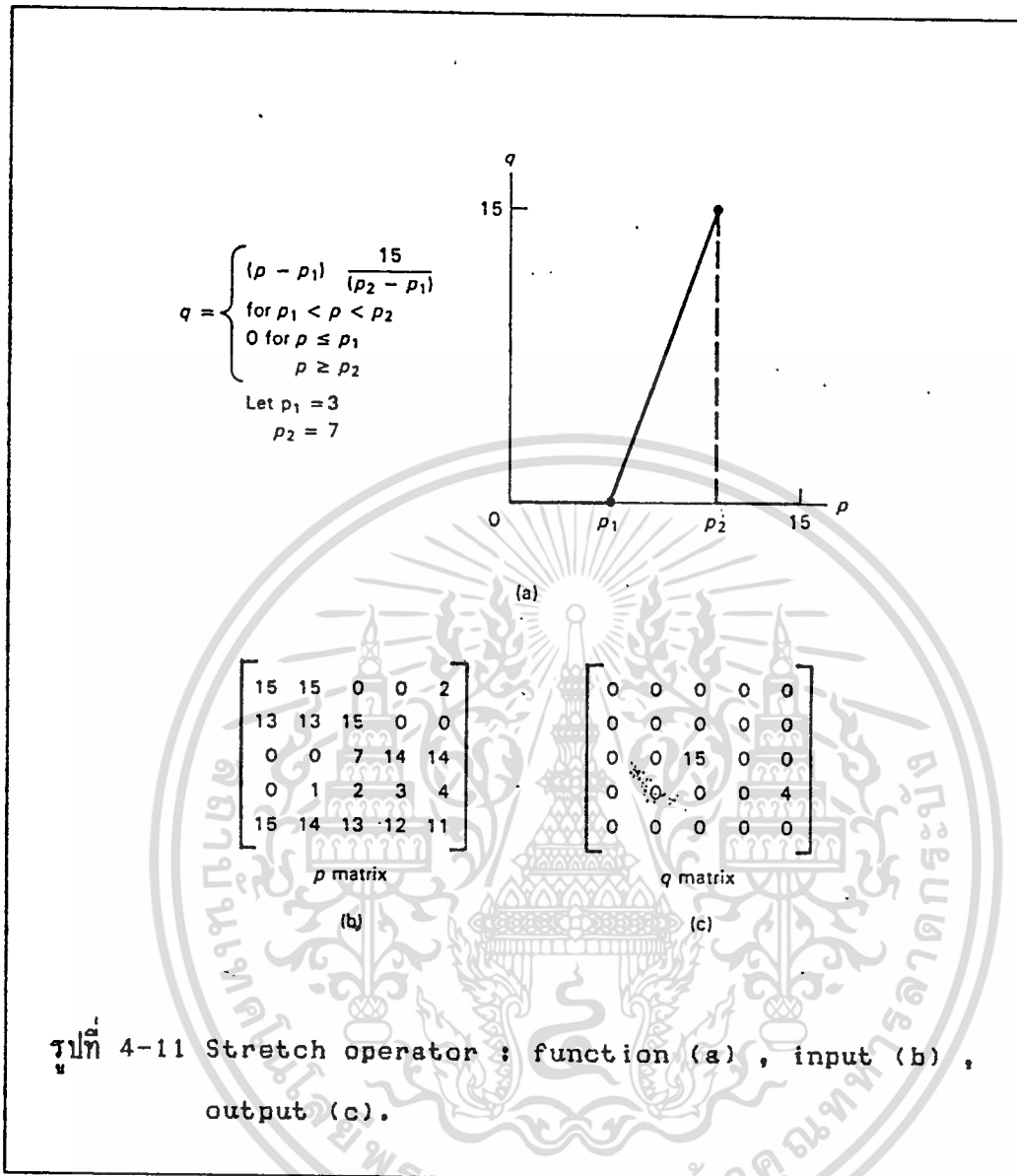


เอาที่หุของสองวิธีเป็นความแตกต่างโดยสมบูรณ์ ผู้ออกแบบระบบหรือผู้วิจัยต้องค้นหาซึ่งวิธีการจัดเตรียมผลที่ได้ทั้งหมดสำหรับรายละเอียดการประยุกต์ใช้งาน

ทั้งสองตัวอย่างแสดงให้เห็นว่า มีหลายทางที่เป็นพื้นฐานการประมวลผลของการบวก การลบ และการคูณ สามารถใช้ปรับข้อมูลใน image matrix ผู้ออกแบบจะต้องคำนวณหาชนิดของข้อมูลที่ต้องการ สำหรับการประยุกต์ใช้งานและเทคนิคการประมวลผลข้อมูล (data processing technique) ซึ่งจะทำให้ลักษณะเด่นของข้อมูลที่ตีขึ้นในข้อมูลภาพ (image data) monadic operator f จะประยุกต์ใช้ $C_{i,j}$ matrix ในรูป 5-16 ค่า $R_{i,j}$ เป็นการปิดเศษสมการ

$$R_{i,j} = (c_{i,j} + 100) \times 255/354$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นอนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



สำหรับการคูณ operator แบบอื่นๆ จะเป็นการสร้าง window เพื่อที่จะลดการคำนวณในบริเวณที่กำลังพิจารณา correction matrix $b_{i,j}$ จะแสดงโดยสมการ

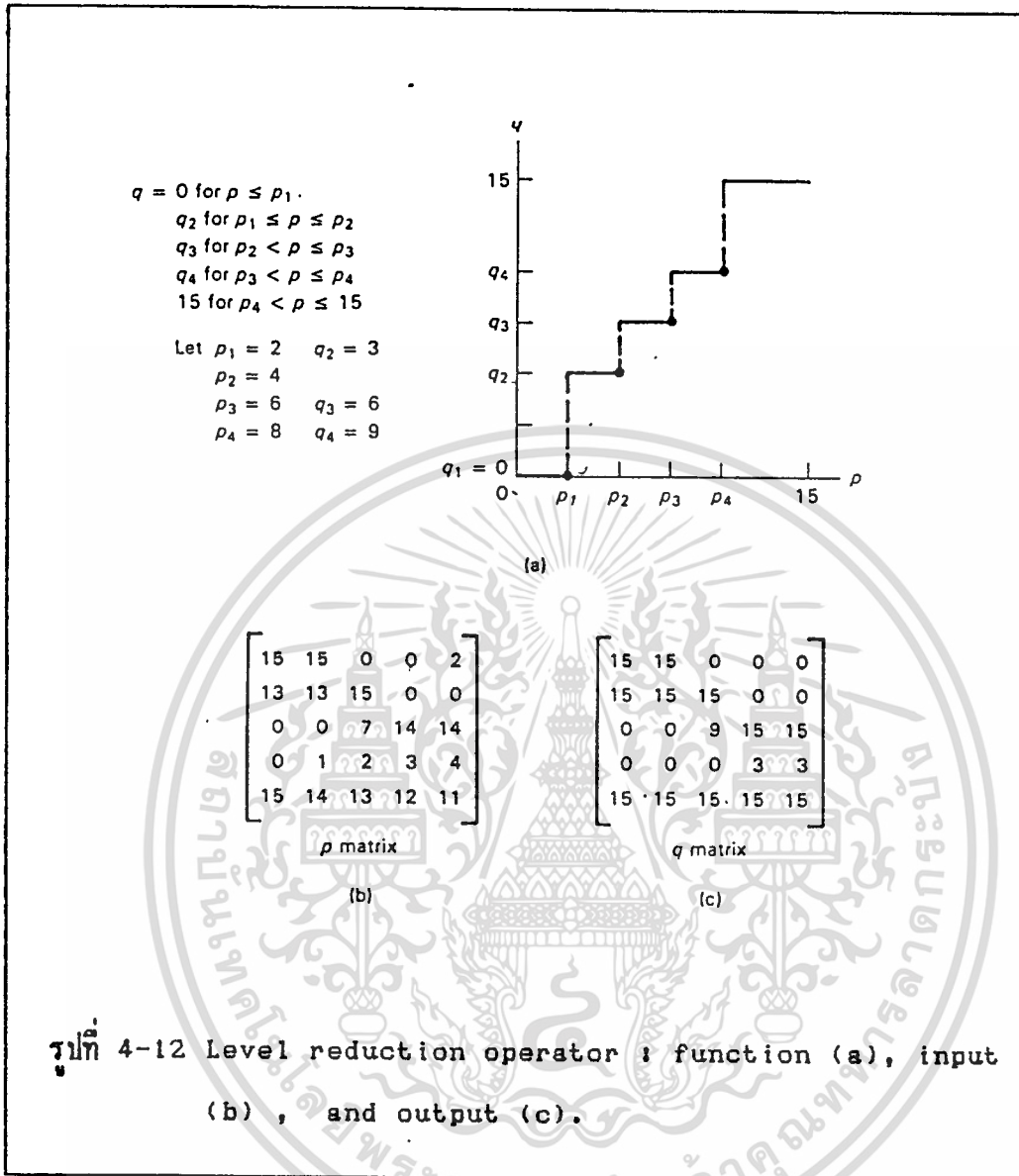
$$c_{i,j} = a_{i,j} + b_{i,j}$$

ซึ่ง $b_{i,j}$ คือ "1" สำหรับทุกค่าที่อยู่ภายในพื้นที่ window และ เป็น "0" สำหรับทุกค่าที่อยู่ภายนอกพื้นที่ window

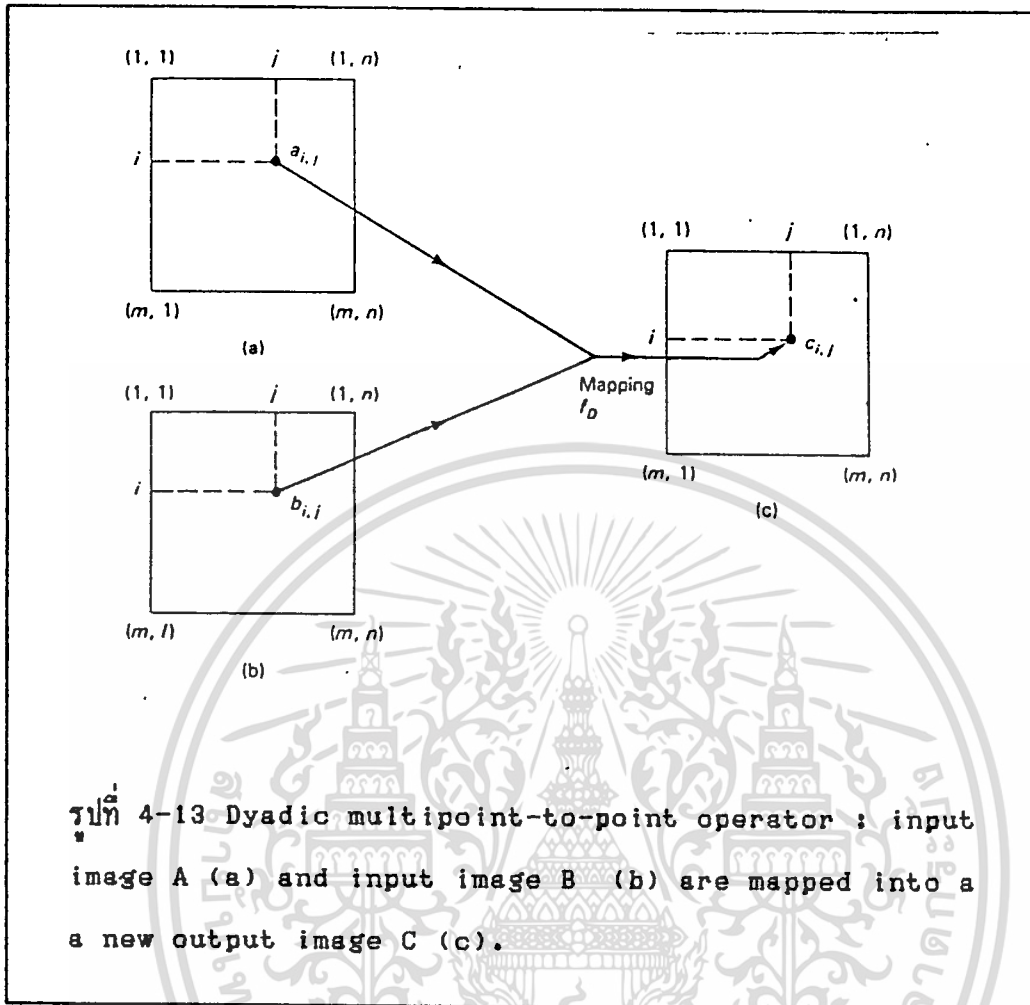
ขนาดของ output image matrix จะมีขนาดเดียวกันกับ input image matrix ไม่มีแถว (row) หรือหลัก (column) หายไปในการประมวลผล แต่จะเกิดขึ้นกับ การ convolution dyadic process สำหรับการ transformation ถูกใช้ในการบวกอิมเมจ ลออิมเมจ และการคูณอิมเมจ ค่า normalization constant อาจจะใช้ระหว่าง dyadic operation

หรือ ฟังก์ชัน monadic scaling operator ใน subsequence phase

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



$\begin{bmatrix} 0 & 12 & 142 & 255 \\ 1 & 6 & 40 & 254 \\ 24 & 0 & 20 & 255 \\ 30 & 2 & 10 & 240 \end{bmatrix}$ <p>$a_{i,j}$ input 1</p>	$\begin{bmatrix} 14 & 11 & 9 & 253 \\ 3 & 5 & 39 & 254 \\ 11 & 1 & 19 & 255 \\ 18 & 2 & 11 & 256 \end{bmatrix}$ <p>$b_{i,j}$ input 2</p>	$\begin{bmatrix} 7 & 12 & 76 & 254 \\ 2 & 6 & 40 & 254 \\ 18 & 1 & 20 & 255 \\ 23 & 2 & 11 & 248 \end{bmatrix}$ <p>$c_{i,j}$ output</p>
--	---	--

รูปที่ 4-14 Image addition for $i=j=4$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(a)

$$\begin{bmatrix} 0 & 12 & 142 & 255 \\ 1 & 6 & 40 & 254 \\ 24 & 0 & 20 & 255 \\ 30 & 2 & 10 & 240 \end{bmatrix} \quad \begin{bmatrix} 14 & 11 & 9 & 253 \\ 3 & 25 & 100 & 0 \\ 11 & 1 & 80 & 1 \\ 30 & 2 & 110 & 255 \end{bmatrix} \quad \begin{bmatrix} -14 & 3 & 133 & 2 \\ -2 & -19 & -60 & 254 \\ 15 & -1 & -60 & 254 \\ 0 & 0 & -100 & -15 \end{bmatrix}$$

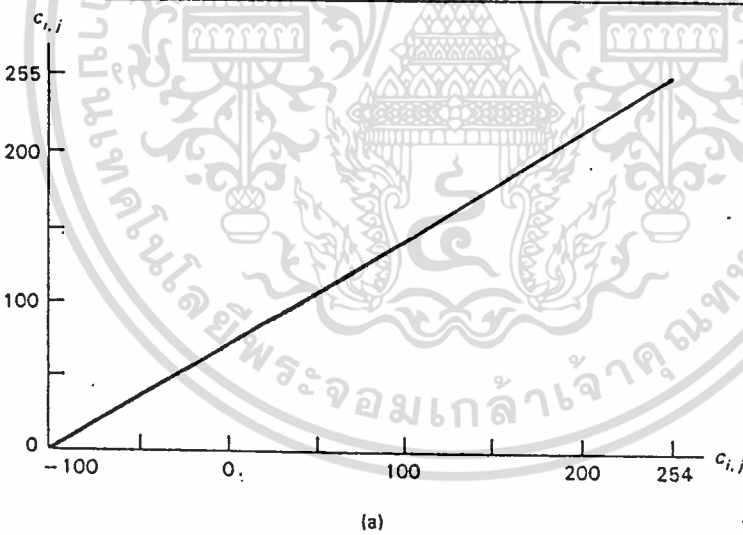
Input #1 = $a_{i,j}$ Input #2 = $b_{i,j}$ Output = $c_{i,j}$

(b)

$$\begin{bmatrix} 14 & 3 & 133 & 2 \\ 2 & 19 & 60 & 254 \\ 15 & 1 & 60 & 254 \\ 0 & 0 & 100 & 15 \end{bmatrix}$$

Output $c_{i,j}$

รูปที่ 4-15 Image subtraction (absolute value) ; simple , but useful only under very limited condition.



$$\begin{bmatrix} 62 & 74 & 168 & 74 \\ 71 & 59 & 29 & 255 \\ 83 & 72 & 29 & 255 \\ 72 & 72 & 0 & 62 \end{bmatrix}$$

Output $c_{i,j}$

(b)

รูปที่ 4-16 Image subtraction with function (a) which generates resultant output (b) having all positive values

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.4 Convolution : spatial transformation

อิมเมจ matrix ใหม่สามารถที่จะถูกสร้างขึ้นได้โดยค่าของพิกเซล ที่ถูกกำหนดให้ในแต่ละตำแหน่งเป็นฟังก์ชันของค่าพิกเซลตำแหน่งข้างเคียง ที่แสดงในรูป 5-18 เป็น 3 x 3 convolution

รูปที่ 4-17 Image multiplication.

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \times \begin{bmatrix} + & + & - \\ - & + & + \\ + & + & - \end{bmatrix} = \begin{bmatrix} | & | & | \\ | & e' & | \\ | & | & | \end{bmatrix}$$

Where:
 $e' = +a + b - c - d + e + f + g + h - i$

รูปที่ 4-18 Convolution (neighborhood) transformation.

ใน 3 x 3 convolution ค่าของพิกเซลของตำแหน่งกลางจะถูกคำนวณบนพื้นฐานค่าที่อยู่ข้างเคียง 9 ตำแหน่ง ตำแหน่งกลางจะถูกเลื่อน (shift) ไปอีกหนึ่งพิกเซล และทำการประมวลผลซ้ำ จนกระทั่งอิมเมจ matrix ถูกสร้างจนเสร็จ มันจะเป็นการแสดงผลของ matrix ผลลัพธ์ ที่ถูกลดขนาดลงสองพิกเซลในแต่ละทิศทางที่เป็นขอบของอิมเมจ ถ้าไม่ต้องการไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ลดขนาดของ matrix สามารถทำได้โดย การเติมข้อมูลที่ เป็นค่าคงที่เข้าไปในแถวและหลักข้าง
เคียงที่อยู่ขอบนอกของอิมเมจ matrix



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

ขอบภาพ

(Edge Enhancement)

5.1 Digital Filter

Digital filter สามารถแบ่งออกได้เป็น low pass filter และ high pass - filter ขึ้นอยู่กับส่วนของ frequency spectrum ที่มีผลกระทบ ผลของสัญญาณรวมจะมีทุกความถี่ ดังแสดงในรูป 5-1

สัญญาณเอาต์พุทของ high pass filter จะมีแต่ความถี่สูง เอาต์พุทของ low pass - filter จะมีแต่ DC และความถี่ต่ำ คุณสมบัติของความถี่ cutoff สามารถเปลี่ยนแปลงโดยเลือกค่าที่จะใช้ในการ filter ได้

ความถี่ต่ำในอิมเมจจะถูกทำให้มีคุณลักษณะ โดยการเปลี่ยนแปลงของความคมชัด (contrast) พิกเซลข้างเคียงอย่างช้าๆ ในความสัมพันธ์กับความถี่สูง จะมีการเปลี่ยนแปลงความคมชัดอย่างรวดเร็วที่บริเวณขอบของวัตถุ

5.2 Low Pass Filter

Low pass filter จะมีผลต่อเฉพาะความถี่ต่ำในข้อมูลภาพ (image data) และจะลดทอนความถี่สูงดังแสดงในรูป 5-2 การลู่มุดที่ไม่ดีในอิมเมจสามารถพิจารณาสัญญาณรบกวนและความถี่สูง เพราะว่าค่าพิกเซลที่อยู่ข้างเคียงมีการเปลี่ยนแปลงอย่างรวดเร็ว ผลของรอยต่างในอิมเมจ สามารถลดลงได้ โดยการใช้ simple averaging filter ล้อมประสิทธิ์ของ matrix สำหรับ averaging filter เป็น 3×3 convolution ซึ่งจะมี 9 element ดังแสดงในรูปข้างล่าง

ค่าของพิกเซลในเอาต์พุท matrix ถูกคำนวณโดยการนำ convolution mask ไป convolution กับ ข้อมูลภาพ (image data) $p_{i,j}$ กับค่าล้อมประสิทธิ์ของ matrix $f_{i,j}$ และให้ผลไว้ที่ตำแหน่งกลางของพื้นที่ในอิมเมจ matrix ใหม่ ค่าใหม่ของ $q_{2,2}$ หาได้จาก

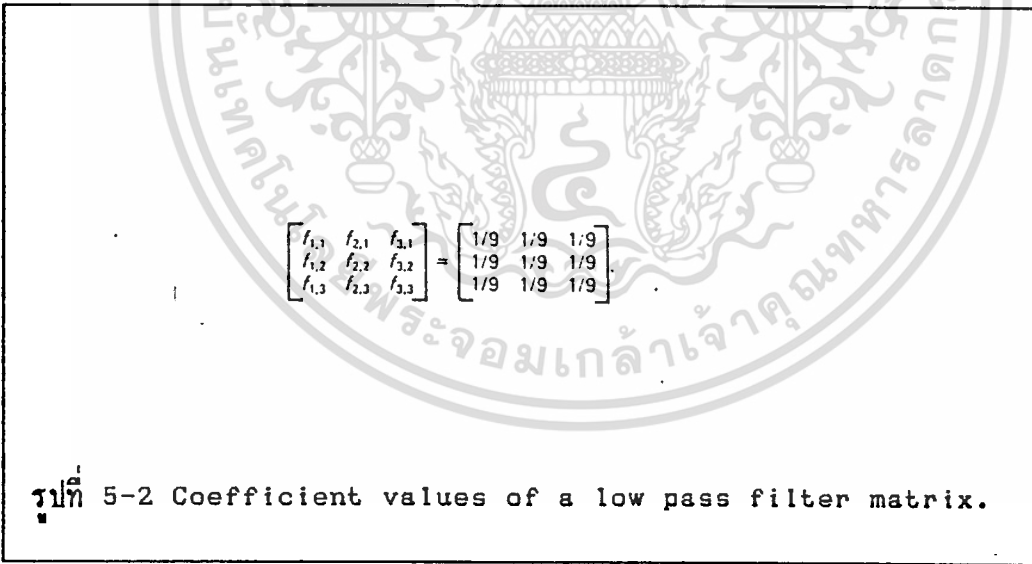
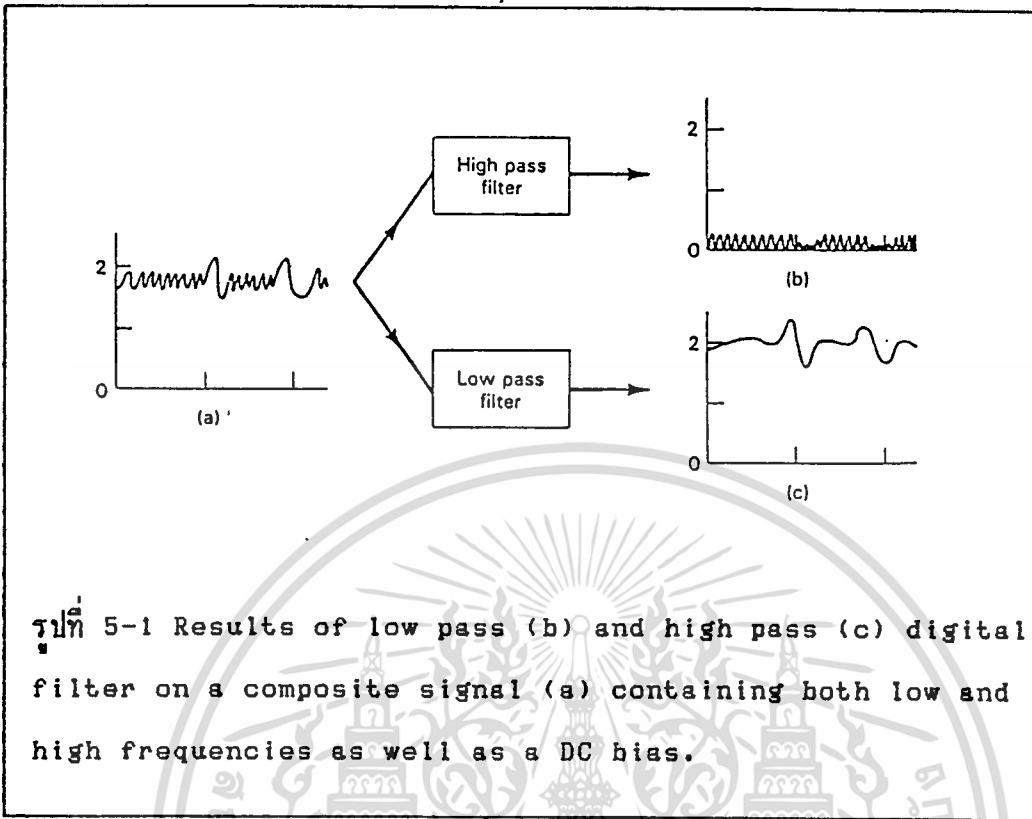
$$q_{2,2} = f_{1,1} \times p_{1,1} + f_{2,1} \times p_{2,1} + \dots + f_{3,3} \times p_{3,3}$$

เงื่อนไขของ low pass filter คือ

(1) ส.ป.ส. ทั้งหมดต้องเป็นบวก และ

(2) ผลรวมทั้งหมดของ ส.ป.ส. จะต้องเท่ากับ "1"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้ภายในเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ถ้าผลรวมของ ส.ป.ส. มากกว่า "1" จะเป็นผลทางการขยาย ถ้าผลรวมน้อยกว่า "1" จะให้ผลทางลดทอน และขนาดของค่าพิกเซลที่อยู่บนอิมเมจจะถูกลดลง ดังนั้น อิมเมจจะมีลดลง การประยุกต์ใช้งานของ ส.ป.ส. ขนาด matrix 3 x 3 กับพื้นที่ข้อมูล ซึ่งทุกพิกเซลมีค่าเหมือนกันหมดจะไม่มีข้อบกพร่อง ซึ่งมี gradient ระหว่างพิกเซลข้างเคียง ผลของ low pass-digital filter ในการลดทอนสัญญาณรบกวน สามารถพิสูจน์ให้เห็นโดยการพิจารณา กรณีที่มีพิกเซลใน matrix ซึ่งเป็น "0" กำหนดให้เป็น bad element ใน sensor array ที่ใช้ ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จับ (acquire) ข้อมูลพิกเซลข้างเคียงจะมีค่าเป็น "high" หมด

$$\begin{bmatrix} 7 & 7 & 7 & 7 \\ 7 & (0) & 7 & 7 \\ 7 & 7 & 7 & 7 \\ 7 & 7 & 7 & 7 \end{bmatrix} = \begin{bmatrix} 7 & 7 & 7 & 7 \\ 7 & (7) & 7 & 7 \\ 7 & 7 & 7 & 7 \\ 7 & 7 & 7 & 7 \end{bmatrix}$$

matrix with
bad element

matrix after
low pass filter

ถ้าค่าของ bad pixel ยังคงเป็นค่า "high" ของ "low" เพราะผลของจุดต่างสีขาว การประยุกต์ใช้งานของ low pass filter จะมีผลกับการลดขนาดของค่าของ bad pixel ผลของการเพิ่มขึ้นของขนาดของบริเวณรอบๆ (neighborhood) เป็น 5 x 5 หรือ 9 x 9 สามารถพิสูจน์โดยการเปรียบเทียบผลของการใช้ low pass filter กับกรณีที่มีพิกเซลเสีย (bad pixel) หนึ่งตัว ("0" ในบริเวณที่อื่นมีค่าเป็น "1" หมด) ถูกแสดงไว้ในตาราง 5-1

TABLE 5-1 Effect of Filter Dimensions

Size of matrix	Convolved Value	Raw Value Bad Element	Relative Computation Time
3 x 3	1/9	8/9	1
5 x 5	1/25	24/25	25/9 = 2-7/9
9 x 9	1/81	80/81	81/9 = 9

พบว่า ขนาดของ matrix ที่ใหญ่จะเป็นเหตุให้ค่าที่คำนวณได้ใหม่ซึ่งเหมาะกับพิกเซลข้างเคียง แต่จะเสียเวลาในการคำนวณ ขนาด 3 x 3 ถูกนำมาใช้ในงานอุตสาหกรรมเป็นส่วนใหญ่ ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เพราะมันลดเวลาในการคำนวณประมวลผล เวลาประมวลผล และ ต้นทุนของเครื่องมือ ผลที่ได้จะถูกตรวจสอบให้แน่นอนว่ามันอยู่ภายใน acceptable error tolerance requirement

5.3 High Pass Filter

High pass digital filter จะกลับกันกับ low pass filter การ filter จะไม่เปลี่ยนแปลงความถี่สูงของสัญญาณ (signal)

ผลของแสงที่เป็น back ground จะมีผลใน DC bias ในระบบการมองเห็นข้อมูลเพราะค่าต่ำสุดของพิกเซลทั้งหมดที่อยู่เหนือบางค่าที่กำหนดให้ ผลของการใช้ high pass filter กับข้อมูลสำหรับอิมเมจขนาด 6×6 ถูกแสดงใน histogram ของรูป 5-3 ค่าพิกเซลที่สูงอาจจะเปลี่ยนแปลงหรือไม่เปลี่ยนแปลงก็ได้ ขึ้นอยู่กับจุดอิ่มตัว (degree of saturation) ของอิมเมจ

ส.ป.ส. ของ high pass filter ตั้งไว้สำหรับ matrix convolution mask ขนาด 3×3 คือ

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

- (1) ส.ป.ส. สามารถเป็นได้ทั้งบวก (positive) หรือ ลบ (negative)
- (2) ผลรวมของ ส.ป.ส. เป็น "0"

เพราะค่ารวมของ ส.ป.ส. เท่ากับ "0" ส่วนที่เป็น DC จะถูกกดไว้อย่างสมบูรณ์ ถ้าผลรวมเท่ากับ "+1" ค่าของ DC จะยังคงมีค่าเท่ากับสัญญาณต้นแบบ (original signal) การ filter ถูกใช้กับ ส.ป.ส. ของ matrix mask ขนาด 3×3 บนข้อมูลภาพ (image data) การคูณ ส.ป.ส. และข้อมูลในทุกค่าของ 9 element และผลบวกทั้ง 9 เพื่อหาค่าใหม่ของตำแหน่ง element ที่จุดกึ่งกลางของ matrix 3×3 mask จะถูกเลื่อนไปอีก "1" และประมวลผลซ้ำ และหาค่าของ element ตัวถัดไป ตัวอย่างที่แสดงการประมวลผลเมื่อหาค่าในรูป 5-4 convolution ของ mask และข้อมูลให้ค่าของตรงกลาง element ที่ตำแหน่ง (2,2)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\begin{bmatrix} +1 & +1 & +1 \\ +1 & -8 & +1 \\ +1 & +1 & +1 \end{bmatrix} \begin{bmatrix} 4 & 8 & 8 \\ 4 & 8 & 8 \\ 4 & 8 & 8 \end{bmatrix} \begin{bmatrix} - & - & - \\ - & [] & - \\ - & - & - \end{bmatrix}$$

mask
data
element

$$\begin{aligned}
 (2,2) &= +(1 \times 4) + (1 \times 8) + (1 \times 8) + (1 \times 4) - (8 \times 8) \\
 &\quad + (1 \times 8) + (1 \times 8) + (1 \times 4) + (1 \times 8) \\
 &= +4 + 8 + 8 + 4 - 64 + 8 + 4 + 8 + 8 \\
 &= +52 - 64 \\
 &= -12
 \end{aligned}$$

ขอบภายในอิมเมจ matrix จะปรากฏการเปลี่ยนแปลงรูปร่างในค่าของพิกเซลข้างเคียง ดังแสดงในรูป 5-4 ตัวอย่าง B ข้อมูลมีความสัมพันธ์กับตำแหน่งและขนาดของขอบภาพซึ่งถูกคำนวณ หรือแทนด้วย high spatial frequency ของค่าพิกเซลในอิมเมจ

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & +9 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

modified coefficient matrix

$$\begin{bmatrix} x & x & x & x & x & x & x \\ x & 4 & 4 & 4 & 4 & 4 & x \\ x & 4 & 4 & 4 & 4 & 4 & x \\ x & 4 & 4 & 4 & 4 & 4 & x \\ x & x & x & x & x & x & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x & x \\ x & 4 & -8 & +20 & 8 & 8 & x \\ x & 4 & -8 & +20 & 8 & 8 & x \\ x & 4 & -8 & +20 & 8 & 8 & x \\ x & x & x & x & x & x & x \end{bmatrix}$$

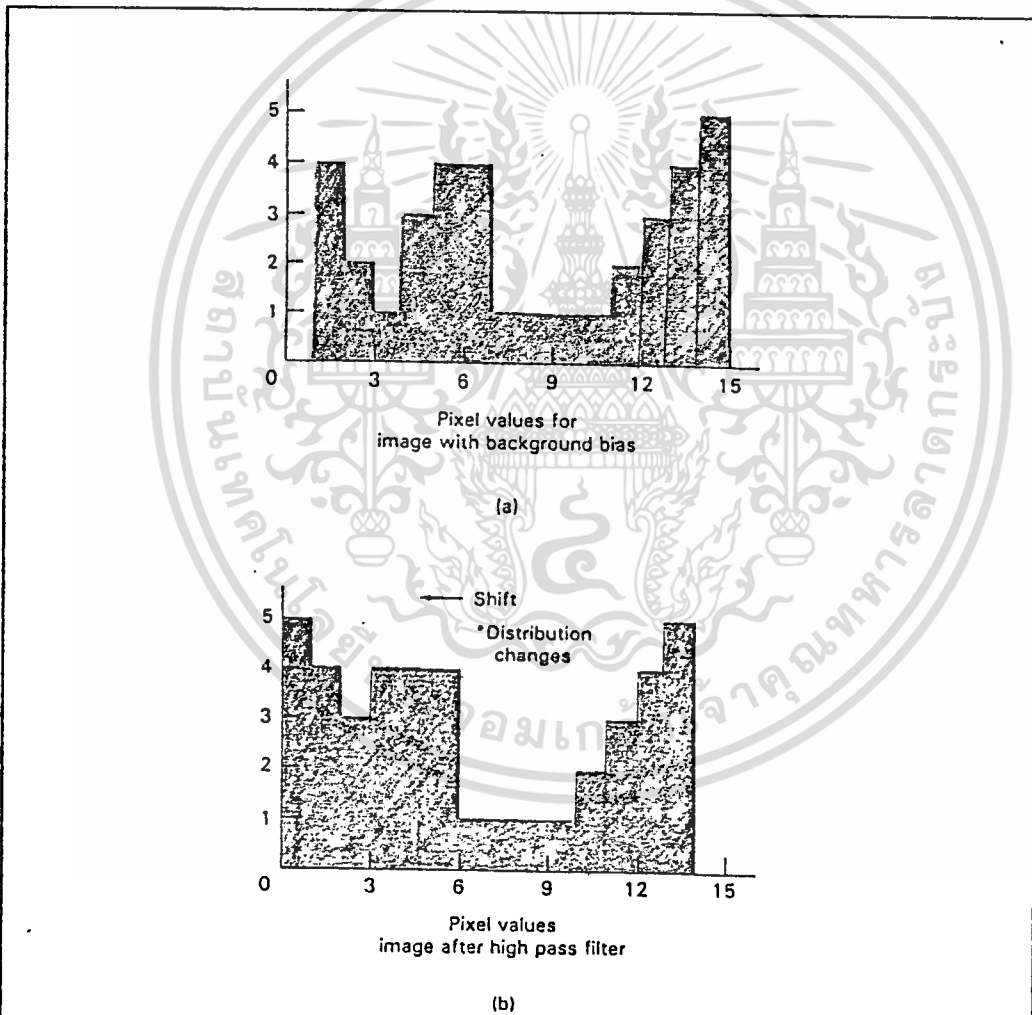
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า effect of modified high pass filter on examples a and b

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

พบว่า high pass filter สามารถกำจัดค่าคงที่ที่ bias ในตัวอย่าง A และทำให้ขอบภาพ (edge pixel) ตีขึ้น ในตัวอย่าง B ซึ่งผลของ matrix มีเส้นของแนว และค่าพิกเซล เป็น "high"

ส.ป.ส.ของ matrix สามารถ modify ซึ่งผลรวมเท่ากับ "1" และผลลัพธ์ที่แตกต่างถูกกำหนดให้ เมื่อ filter ถูกนำไปใช้กับข้อมูลในตัวอย่าง A และ B

ผลของการ convolution ของการ modify high pass filter กับข้อมูลในอิมเมจ A และ B จัดเตรียมอิมเมจไว้ให้ ซึ่ง bias ก็ยังคงอยู่ และความถี่สูงจะถูกขยาย (amplified)

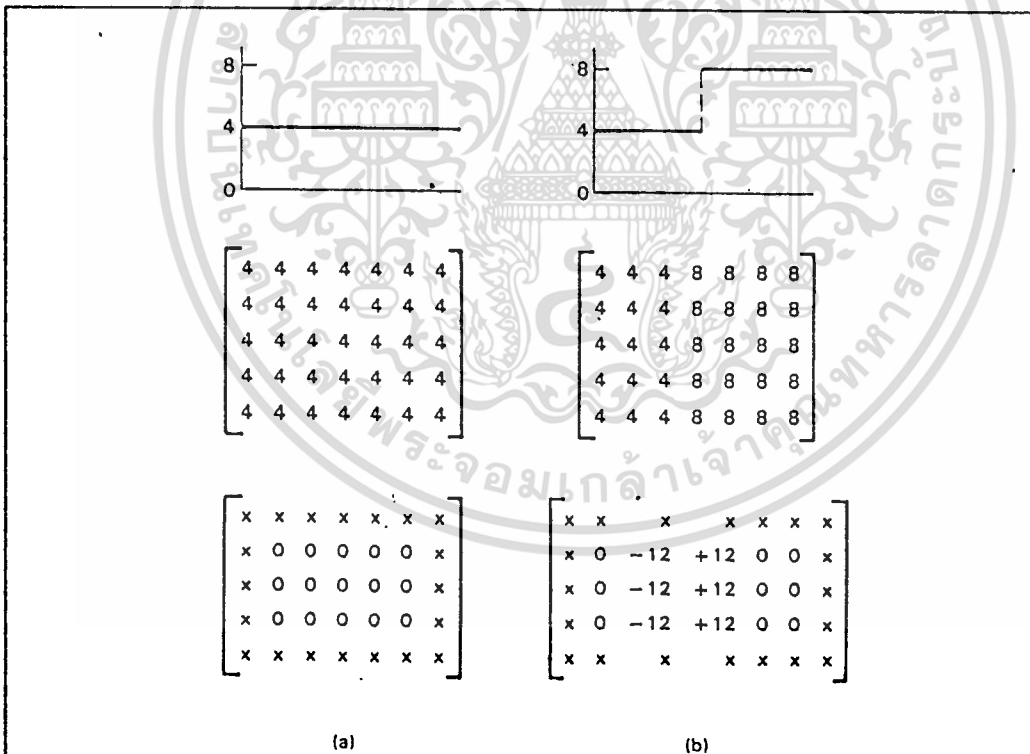


รูปที่ 5-3 Effect of a high pass filter on a histogram. When values of pixels are shifted, the shape of the histogram changes.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

matrix ขนาด 3 x 3 เป็นส่วนที่เล็กที่สุดและต่ำที่สุดในการคำนวณที่ต้องทำการ convolution วิธีที่ง่าย คือ ขยายเป็น 5 x 5 และ 9 x 9 matrices แต่ผลอาจจะให้ความแตกต่างไม่เด่นชัด ขณะที่การคำนวณขยายโดยองค์ประกอบ (factor) ของ 2.7 และ 9 นั้น การ convolution ที่มีขนาด 3 x 3 สามารถทำให้มีประสิทธิภาพสำหรับการใช้งานในงานอุตสาหกรรม ผลของ 9 x 9 สามารถทำบนระบบ 3 x 3 โดยการซ้ำกับการ convolution ขนาด 3 x 3 ค่าที่อยู่นอกเหนือส่วนนี้จะสูญหาย เพราะว่าเป็นข้อมูลที่ไม่สำคัญ high pass filter มีคุณสมบัติ ซึ่งข้อมูลในบริเวณขอบถูก modify และขอบมีผลต่อการขยาย edge detection เป็นความรู้พื้นฐานในการหาจากอิมเมจ เพราะว่าข้อมูลมีความสัมพันธ์กับ contrast , chape , location และ dimemsion

ขอบภาพ (edge) เป็นพารามิเตอร์วิกฤต (critical parameter) ของแนวความคิด และตำแหน่งของขอบภาพ เป็นสิ่งสำคัญยิ่ง ในระบบการมองเห็นจะใช้ความแตกต่างจาก high pass filter ไปทำให้จุดเด่น (feature) ที่กำลังพิจารณาอยู่ให้ดีขึ้น



รูปที่ 5-4 Results of high pass filter on image not having an edge (a) and on an image containing an edge (b). x equals last value because of incomplete data.

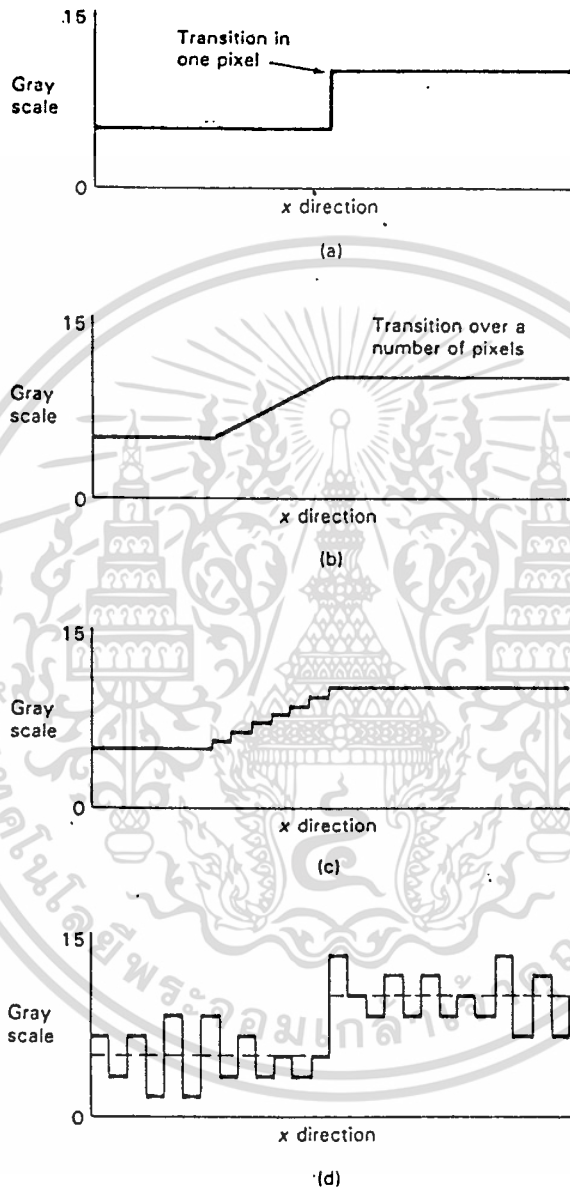
5.4 Edge Enhancement

ขอบภาพคือบริเวณรอบๆ ระหว่างสองบริเวณที่มีความแตกต่างของ gray level ขอบภาพในอุดมคติจะเป็นฟังก์ชันแบบขั้น (step function) และสามารถที่จะเห็นความคมชัดหรือพรมัว ซึ่งความพรมัวหรือขอบภาพที่หนาสามารถพิจารณาองค์ประกอบของจำนวนตำแหน่งขอบภาพที่เล็ก ดังในรูป 5-5 ในความเป็นจริงของการใช้งานระบบมองเห็น ขอบภาพโดยปกติจะมีสัญญาณรบกวน (noise) รวมอยู่ในข้อมูลด้วย

ความถี่สูงก็ถูกรวมอยู่กับการเปลี่ยนแปลงที่ชัดเจนในค่าความเข้มแสงของพิกเซล การทำให้ขอบภาพดีขึ้นสามารถทำได้โดย การใช้รูปแบบของ high pass filter กับข้อมูลภาพ (image data) ตัว operator ที่มีขนาดใหญ่ จะทำให้ขอบภาพดีขึ้นที่ยังคงรูปร่างของตัวอักษร แต่ที่นิยมใช้ในงานอุตสาหกรรม คือ

- Laplacian edge enhancement
- Robert's gradient
- Sobel edge detector

ขอบภาพแบบขั้น (step) จะแยกพื้นที่ออกเป็นสองส่วนกับความสัมพันธ์กับความสม่ำเสมอของ gray level บนแต่ละข้างของขอบภาพ เส้นที่พิจารณาจะมีขอบภาพสองขั้น (2 step) คือ อันหนึ่งเป็นบวก และอีกอันหนึ่งเป็นลบ ความกว้างของเส้นขึ้นอยู่กับการแยกระยะทางระหว่างสอง step



รูปที่ 5-5 Types of edges in an image : ideal step (a), wide (b), break down of wide edge (c), and ideal step with noise (d).

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.4.1 Laplacian Edge Enhancement Operator

Laplacian operator ที่ใช้ในระบบการประมวลผลเป็นพื้นฐานทางคณิตศาสตร์ อนุพันธ์ย่อยของ laplacian แสดงในรูปของฟังก์ชันต่อเนื่องได้ดังนี้

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

สมการคณิตศาสตร์ที่ต่อเนื่อง สามารถประมาณโดยความแตกต่างของ operator ในการ match ที่ไม่ต่อเนื่อง ของการใช้งานประมวลผลการมองเห็นแบบดิจิทัล Rotationally insensitive ของการประมวลผล laplacian จำเป็นที่จะต้องหาการเปลี่ยนแปลงความคมชัดของความเข้มแสงที่พิกเซล ในทิศทาง x และ y

$$L(i, j) = \nabla^2 p(i, j) = \Delta x^2 p(i, j) + \Delta y^2 p(i, j)$$

เมื่อ

$$\Delta x^2 = [p(i-1, j) - p(i, j)] - [p(i, j) - p(i+1, j)]$$

$$\Delta y^2 = [p(i, j+1) - p(i, j)] - [p(i, j) - p(i, j-1)]$$

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

$$\Delta x^2 = [d-e] - [e-f]$$

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

$$\Delta y^2 = [h-e] - [e-b]$$

เมื่อรวมพจน์เข้าด้วยกันจะได้ผลดังข้างล่างซึ่งเป็น Laplacian operator

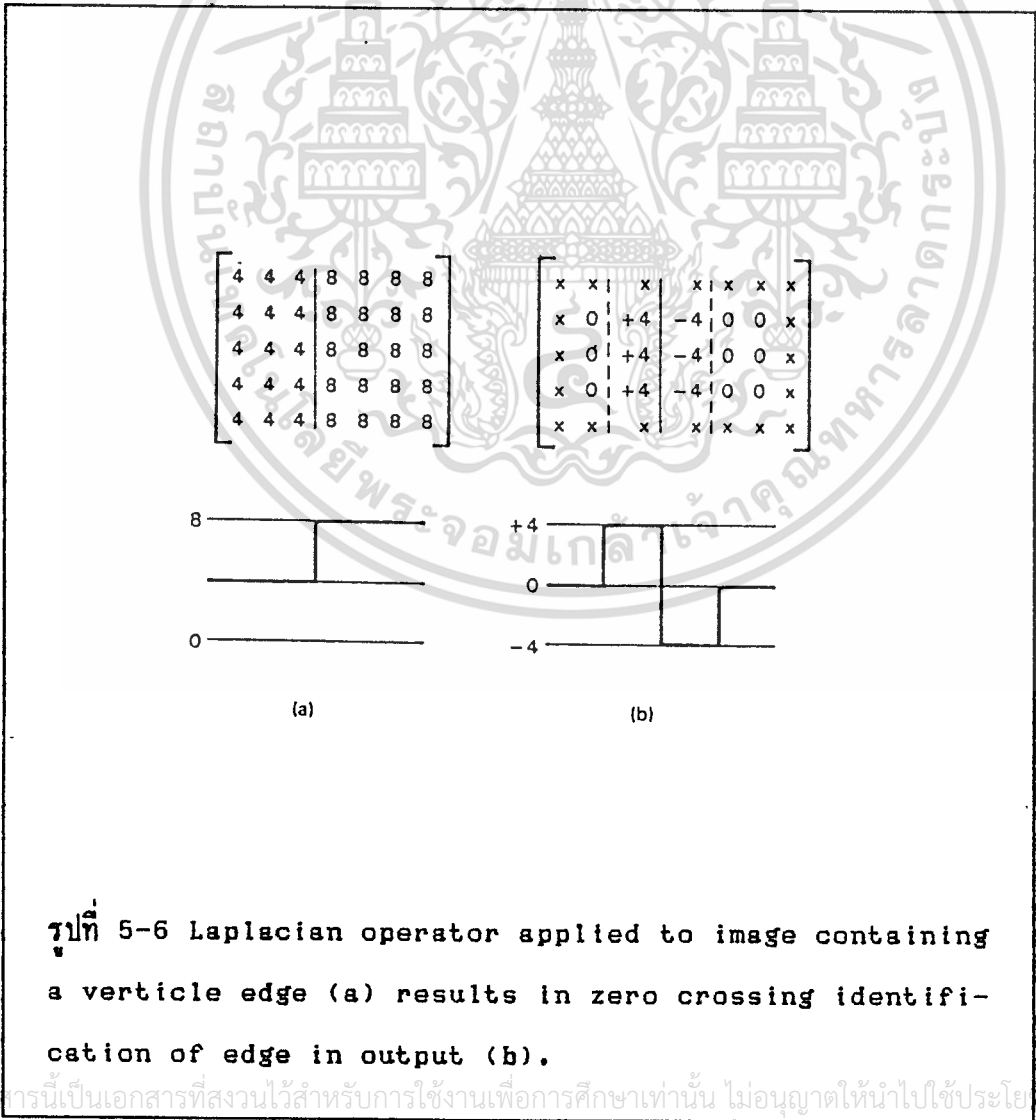
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

$$L(i, j) = b + d + f + h - 4e$$
 ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$L(i, j)$ สามารถลดรูปให้เป็น mask ได้ดังนี้

$$\begin{bmatrix} 0 & +1 & 0 \\ +1 & -4 & +1 \\ 0 & +1 & 0 \end{bmatrix}$$

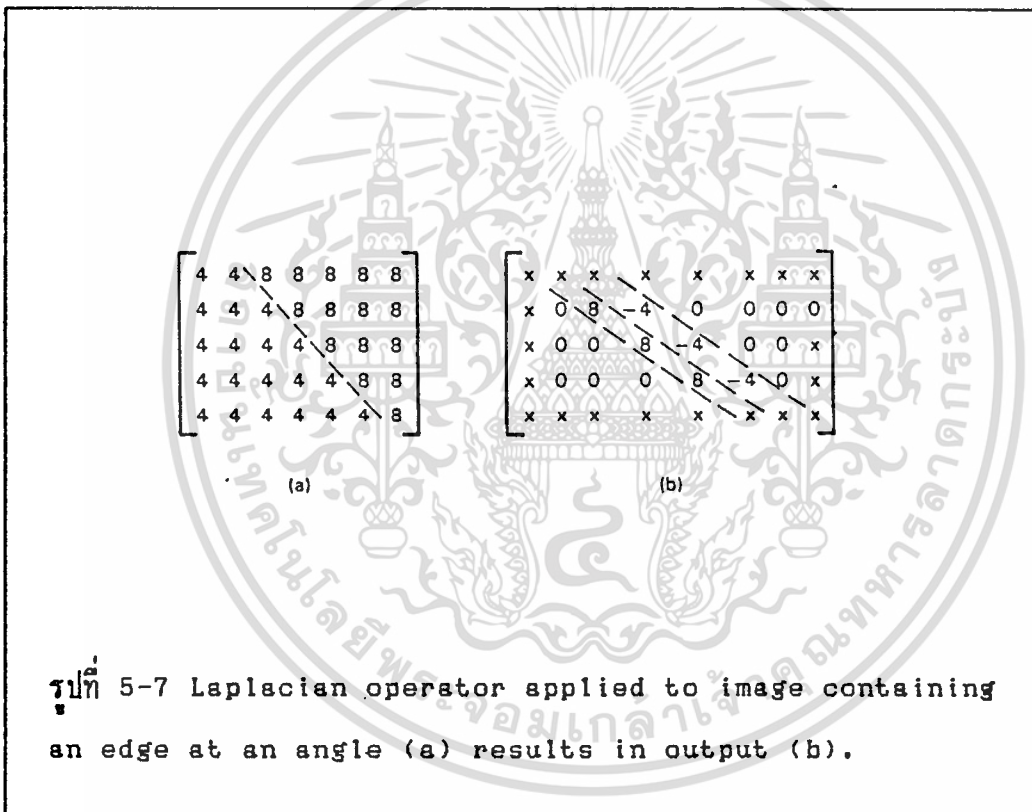
สรุป Laplacian operator คำนวณหาความแตกต่างระหว่าง gray level ของจุดศูนย์กลางของพิกเซล และค่าเฉลี่ยของ gray level ของพิกเซลข้างเคียง 4 ตัว ในทิศทางแนวนอนและแนวตั้ง Laplacian operator เป็น high pass filter เพราะว่าผลรวมของส.ป.ส. เป็น "0" และมันมีทั้ง ส.ป.ส. ที่มีค่าเป็นบวกและลบ การใช้งานของ Laplacian กับ input matrix ของตัวอย่าง B ประกอบด้วยขอบภาพทางแนวตั้ง ดังแสดงในรูป 5-6



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

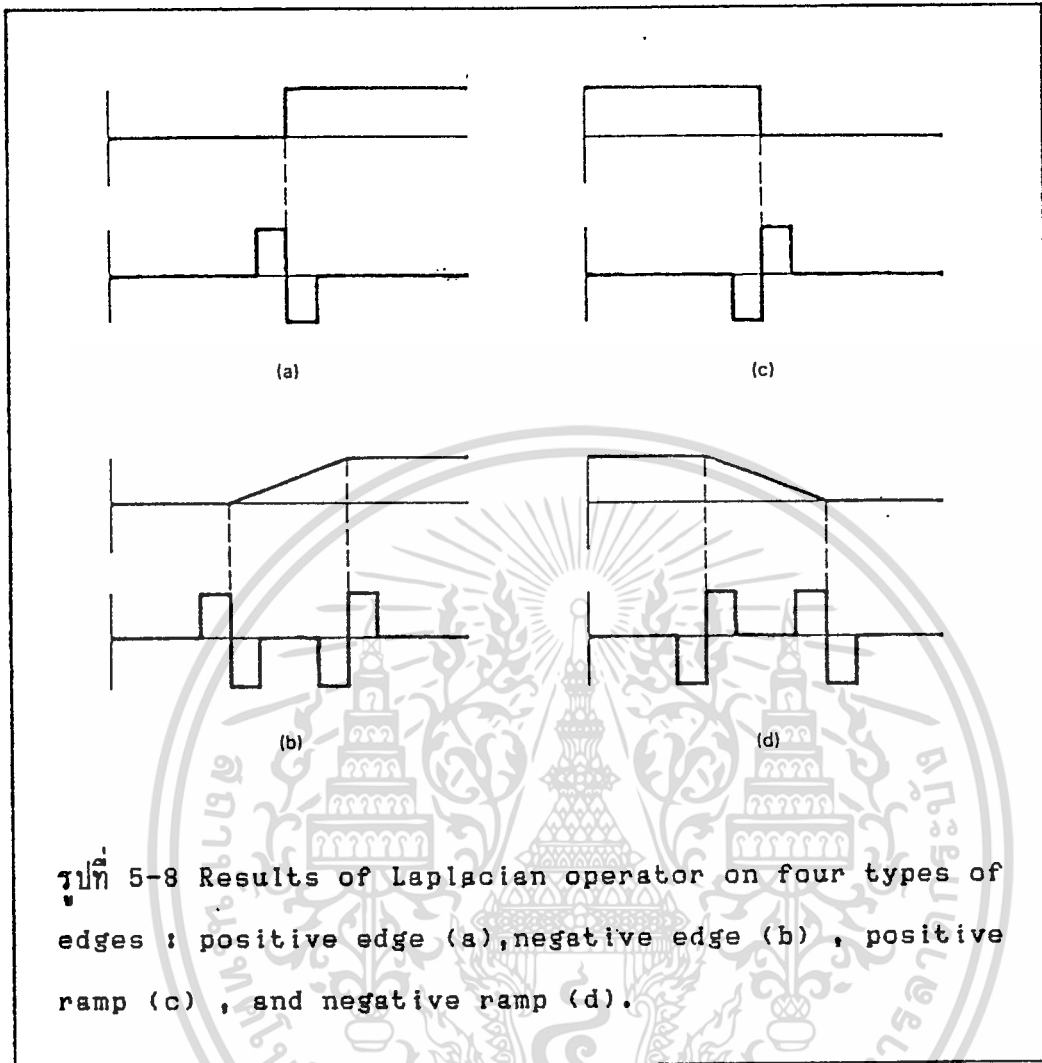
การใช้งานของ Laplacian กับข้อมูลภาพและขอบภาพที่มีมุมดังรูป 5-7

การแสดงของขอบภาพที่เป็นบวกที่แสดงในรูป 5-8 มีผลใน positive impulse ตามด้วย negative impulse เพราะ Laplacian operator กำลังทำอนุพันธ์กับผิว ขอบภาพจะถูกกำหนดตำแหน่งระหว่างสอง impulse ฟังก์ชันที่จุดซึ่ง impulse function ตัดผ่าน "ZERO" ดังนั้น Laplacian operator สามารถใช้ทำให้ขอบภาพดีขึ้นได้ และการตรวจจับการตัดผ่าน "ZERO" สามารถใช้กับตำแหน่งของขอบภาพ impulse function จาก 4 รูปแบบของขอบภาพดังแสดงในรูป 5-8



ขอบภาพและเส้นในขอบภาพ สามารถทำงานได้ในทิศทางต่างๆ มันเป็นความต้องการอย่างสูง ซึ่ง operator เป็น Isotropic ถ้ามันถูกใช้พัฒนาอิมเมจที่มีจุดสนใจไม่ชัดเจนในอิมเมจเพื่อที่จะให้เป็น Isotropic ตัว operator จะต้องเกี่ยวข้องกับเฉพาะอนุพันธ์ที่มี order เป็นเลขคู่ ผลการคำนวณจากการใช้ Isotropic operator กับอิมเมจและการหมุนเอาที่พุก จะเป็นเอกลักษณ์ (identical) กับการคำนวณ เมื่ออินพุตถูกหมุนและเมื่อ operator ถูกนำมาใช้ operator ที่ใช้ในการพัฒนาขอบภาพโดยปกติเกี่ยวข้องกับผลต่างกับ identity ของการเปลี่ยนแปลงความชันของค่ารวมกับค่าของพิกเซลเหมือนบริเวณผิว อนุพันธ์อันดับที่ 2 อาจจะเป็นเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ประมวลผลโดยความแตกต่างอันดับหนึ่ง

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



$$\frac{\partial^2 f}{\partial x^2} = (f_{m,n} - f_{m+1,n}) - (f_{m+1,n} - f_{m+2,n})$$

$$\frac{\partial^2 f}{\partial x^2} = f_{m,n} - 2f_{m+1,n} + f_{m+2,n}$$

อินพุต a, b, c การประมาณของอนุพันธ์อันดับสองในทิศทาง x

$$\frac{\partial^2 f}{\partial x^2} = a - 2b + c$$

$$\begin{bmatrix} \text{Mask} \\ | -2 | \end{bmatrix}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การแสดง matrix ซึ่งจะใช้สำหรับ $f_{i,j}$ คือ

row : $i = 1 \longrightarrow M$ (y direction)
 Column : $j = 1 \longrightarrow N$ (x direction)

MN matrix มี M แถว และ N หลัก

$$\begin{bmatrix} f_{1,1} & \dots & f_{1,N} \\ \vdots & & \vdots \\ f_{M,1} & \dots & f_{M,N} \end{bmatrix}$$

5.4.2 Robert's Gradient Operator

Robert's gradient operator (รูป 5-9) เป็นวิธีที่ง่ายกว่า Laplacian - operator เพราะว่ามีกระทำการบนพื้นที่ 2×2 ของพิกเซลของแต่ละจุด การใช้งานของข้อมูลจากขนาดพื้นที่น้อยกว่าพื้นที่ 3×3 ของ Laplacian operator จะลดเวลาในการคำนวณลง และความสามารถของระบบ operator ใช้กับอนุพันธ์ในแนวทแยงเพื่อประมาณค่า gradient ที่เป็นจุดขนาดของ operator เท่ากับรากที่สองของผลรวมของผลต่างยกกำลังสอง มันสามารถประมาณโดยใช้ สมการคำนวณแบบง่าย ๆ ของผลรวมของค่าสมบูรณ์ของแต่ละผลต่างในแนวทแยง

Magnitude operator = $(\Delta 1^2 + \Delta 2^2)^{1/2}$

Absolute value estimate = $|\Delta 1| + |\Delta 2|$

where $\Delta 1 = p(i, j) - p(i+1, j+1)$ and $\Delta 2 = p(i+1, j) - p(i, j+1)$

a	b
c	d

New $q(1,1) = |\Delta 1| + |\Delta 2|$

where $\Delta 1 = (a-d)$ and $\Delta 2 = (b-c)$

2	5	6	2
3	1	3	4
2	3	5	1

input

New $q(1,1) = (2-1) + (5-3) = |1| + |2| = 3$

3	7	3	x
1	4	3	x
x	x	x	x

outputs after
Robert's
Operator

รูปที่ 5-9 Robert's gradient operator.

5.4.3 Sobel Edge Detector Operator

Sobel operator (ในรูป 5-10) เป็นการคำนวณแบบไม่เป็นเชิงเส้นของขนาดขอบภาพที่เป็นจุด แต่มันจะไม่ใช้ค่าที่จุดของมันเองในการคำนวณ ค่าของนิคเซล กำหนดให้โดย

$S = (\Delta x^2 + \Delta y^2)^{1/2}$

where

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \cdot \Delta x = (a + 2d + g) - (c + 2f + e)$$

$$\Delta y = (g + 2h + i) - (a + 2b + c)$$

input

$$\Delta x = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \quad \Delta y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

convolution masks for sobel operator

3	4	2	5	1	x	x	x	x	x
2	①	6	4	2	x	q _{2,2}			
3	5	7	1	3					
4	2	5	7	1					
2	5	1	3	2					

input

$q_{2,2} = (\Delta x^2 + \Delta y^2)^{1/2}$ around pixel 2,2

$$\Delta x = 3 + 4 + 3 - 2 - 12 - 7 = +11$$

$$\Delta y = -3 - 8 - 2 + 3 + 10 + 7 = +17$$

$$q_{2,2} = [(-11)^2 + (+7)^2]^{1/2}$$

$$= [121 + 49]^{1/2}$$

$$= (170)^{1/2} \approx 13$$

รูปที่ 5-10 Sobel edge detector operator.

5.4.4 Other Local Operator

วิธีการพัฒนาทนในส่วนที่กล่าวมาก่อนหน้านั้นเป็นแบบพื้นๆ และพบใน standard processing package ของโปรแกรมย่อย (software routines) ใช้กับระบบมองเห็นในทางตัวเลข Convolution mask รูปแบบต่างๆ ดังแสดงไว้ให้ดูข้างนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Convolution mask

CONVOLUTION MASKS

1. Low pass filters (coefficients positive and added to one)

$\begin{bmatrix} 0.1 & 0.1 & 0.1 \\ 0.1 & 0.2 & 0.1 \\ 0.1 & 0.1 & 0.1 \end{bmatrix}$	$\begin{bmatrix} 1/16 & 1/8 & 1/16 \\ 1/8 & 1/4 & 1/8 \\ 1/16 & 1/8 & 1/16 \end{bmatrix}$
---	---

2. High pass filters (coefficients added to zero)

$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}$
---	---

3. High pass filter with DC Bias (coefficients added to one)

$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{bmatrix}$	$\begin{bmatrix} 1 & -2 & 1 \\ -2 & 5 & -2 \\ 1 & -2 & 1 \end{bmatrix}$
---	---

4. Gradient-directional (coefficients added to zero)

$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ -1 & -1 & -1 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 1 \\ -1 & -2 & 1 \\ -1 & -1 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & 1 & 1 \\ -1 & -2 & 1 \\ -1 & 1 & 1 \end{bmatrix}$
N	NE	E

5. Shift and difference filters (coefficients added to zero)

$\begin{bmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$
Vertical edges	Horizontal edges	Horizontal and vertical edges

6. Blurs

$\begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
Horizontal	Vertical	Diagonal

7. Difference filters

$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}$
Vertical	Horizontal

8. Vertical differentiation filters

Take absolute value difference (7).

9. Horizontal differencing and vertical smoothing filters

$\begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}$
--

10. Laplacian filters (coefficients added to zero)

$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}$
---	---

11. Bright region expansion

Maximum of

$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

Convolution mask

12. Medium filter
(reduce camera noise)

Fifth largest of $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

13. Enhance line segment
(coefficients added to zero)

Vertical $\begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix}$ Horizontal $\begin{bmatrix} -1 & -1 & -1 \\ 2 & -2 & 2 \\ -1 & -1 & -1 \end{bmatrix}$ L-R Diagonal $\begin{bmatrix} -1 & -1 & 2 \\ -1 & 2 & -1 \\ 2 & -1 & -1 \end{bmatrix}$

14. Edge detector

$\text{Max}(A) - (B)$
where: $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ $B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

15. Line detector

$G ::= A - \text{Min}(\text{Min}(\text{Max}(\text{Max}(B))))$
where: $A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ $B = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

16. Template matching
(Cross correlation)

$\begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

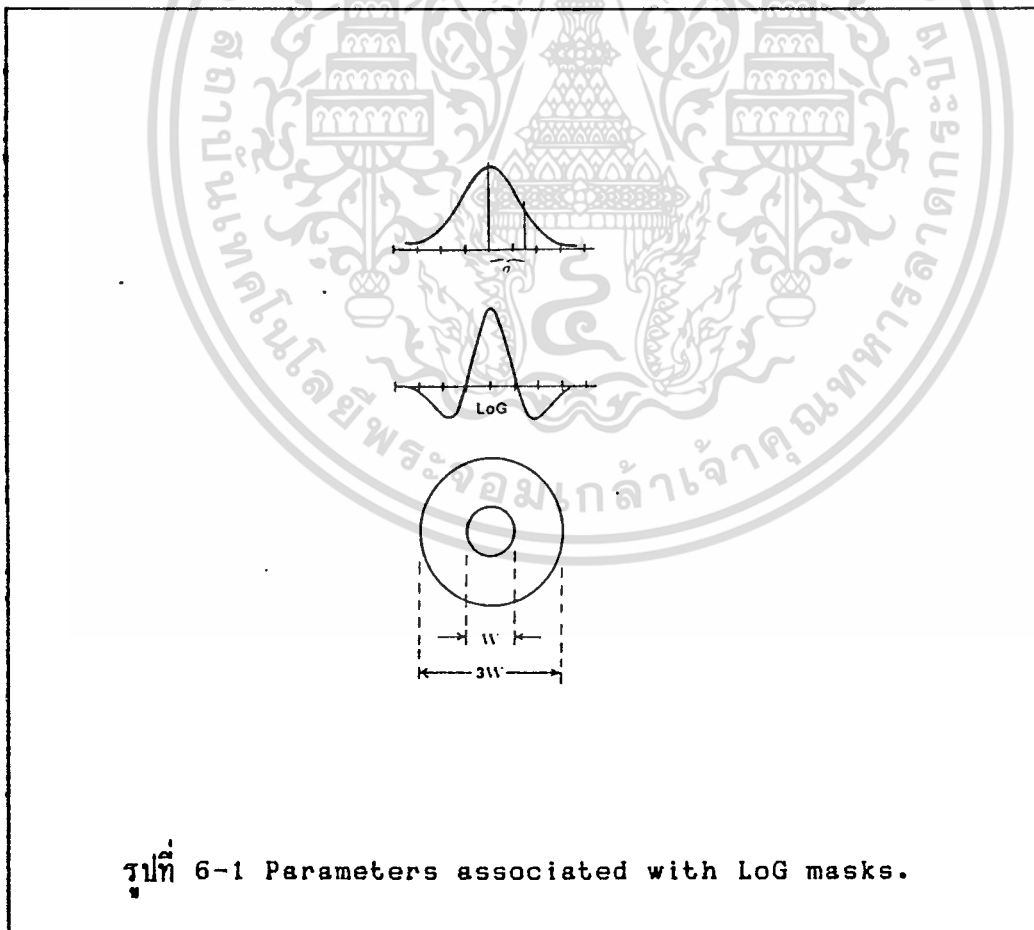
บทที่ 6
การหาซีโร-ครอสซิง
 (Zero-crossing)

6.1 ZERO-CROSSING OF LAPLACIAN-GAUSSIAN FILTER IMAGE

สูตรที่ใช้ในการหา convolution mask [โดย Marr] ใน 2 มิติ คือ

$$\nabla^2 G(x, y) = \frac{1}{2\pi\sigma^4} \left(2 - \frac{(x^2 + y^2)}{\sigma^2} \right) \cdot \exp \left[-\frac{(x^2 + y^2)}{2\sigma^2} \right] \quad (1)$$

ซึ่ง σ เป็นค่าคงที่ของ Gaussian และเรากำหนดให้ $w = 2\sqrt{2}$ แทนความกว้างของ central excitatory region ของตัว operator ซึ่งค่า parameter ต่างๆได้แสดงไว้ในรูปที่ 6-1

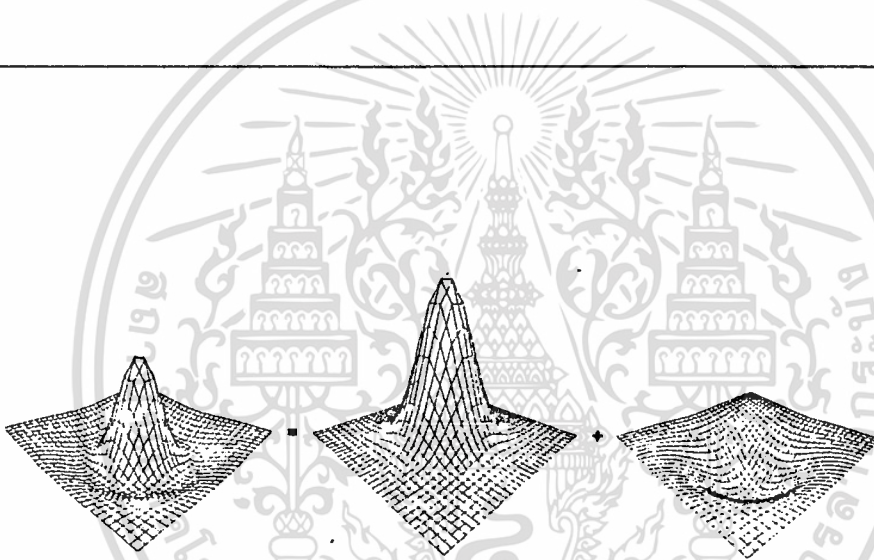


รูปที่ 6-1 Parameters associated with LoG masks.

ตามทฤษฎีของ Marr-Hildreth ที่ว่าด้วยการตรวจหาขอบของภาพ (edge detection) ซึ่งได้กล่าวไว้ว่า ความเข้มของขอบภาพ (intensity edge) เกิดขึ้นจาก ความละเอียดที่แตกไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ต่างกัน (different resolution) จะถูกพบโดย convolution ภาพกับอนุกรมของ LoG filter ที่มีความแตกต่างของ spatial frequency parameters และจากนั้นหาเส้นทางซึ่ง convolution แล้วให้ output ตัดผ่าน "ZERO"

การเปลี่ยนแปลงของความเข้ม (intensity) ถูกตรวจพบในภาพอิมเมจ $I(x,y)$ โดยการหา zero-crossing ในรูปสมการ $\nabla^2 G(x,y) * I(x,y)$ เมื่อ * แสดงถึงการ convolution ความไวในการกรอง (filter sensitivity) ต่อการเปลี่ยนแปลงความเข้ม ขึ้นอยู่กับค่า W (W = ความกว้างของ central excitatory region ของ operator) และค่าคงที่ Gaussian (σ) ซึ่งถ้ามีค่าน้อยๆ จะมีความไวในการ filter

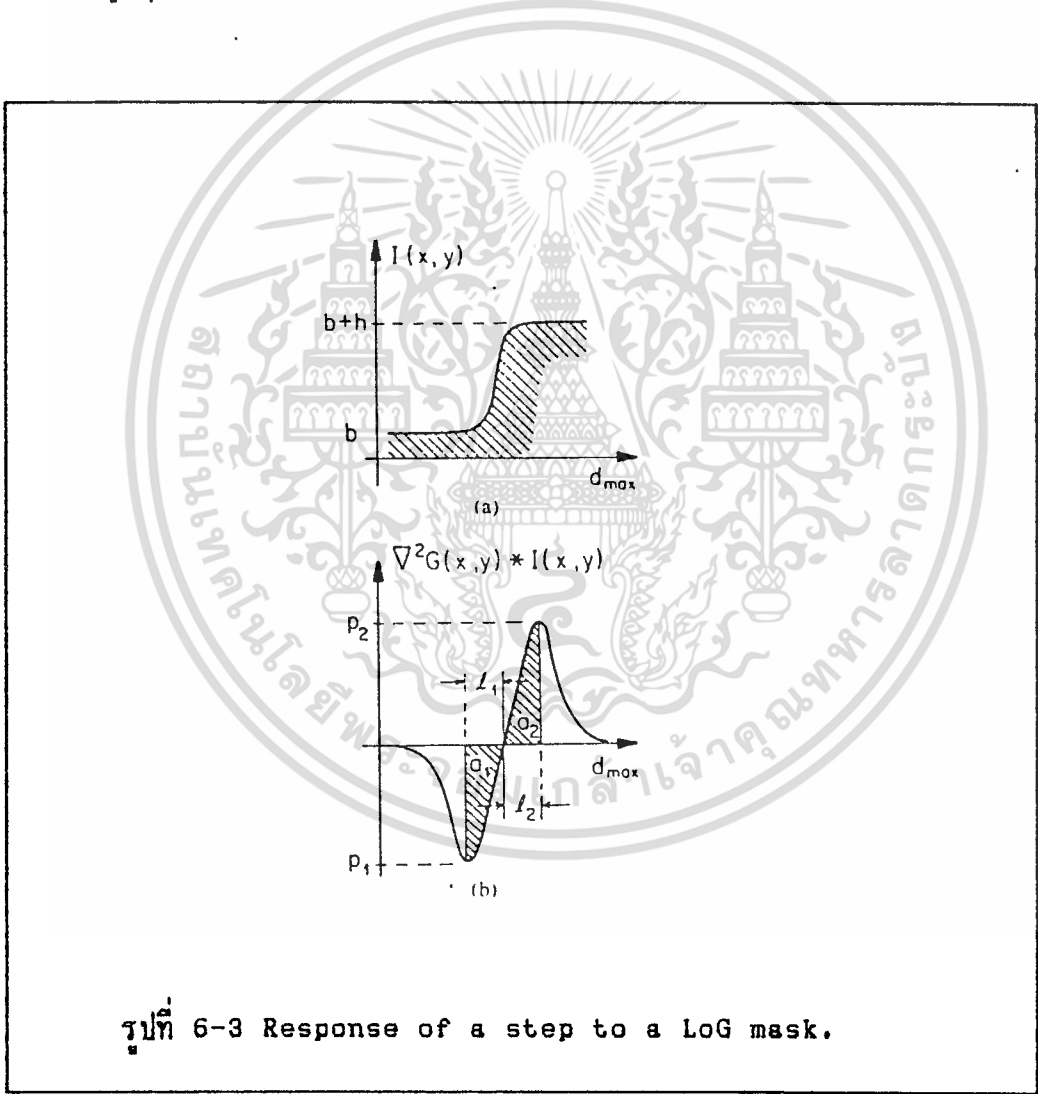


รูปที่ 6-2 The point-spread function corresponding to Gaussian filtering followed by differencing is closely approximated by the difference of two appropriate Gaussians.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6.2 การตรวจจับการเปลี่ยนแปลงความเข้ม (Detection of Intensity change)

เมื่อมีการเปลี่ยนแปลงความเข้มเกิดขึ้น จะทำให้เกิดจุดยอด (peak) ใน first directional derivative ของความเข้ม และ zero-crossing ใน second directional derivative ดังรูปที่ 6-3 การตรวจจับการเปลี่ยนแปลงของความเข้มถูกทำให้ลดลง เพื่อที่จะหา zero-crossing ใน second derivative ของความเข้มในทิศทางของความชันสูงสุด (maximum slope) ณ ตำแหน่ง zero-crossing โดยใช้ orientation-independent operator จำนวนการทำ convolution จะถูกลดลงเหลือเพียง 1 ครั้ง และ ทิศทางหาได้จากค่าความชันสูงสุด



รูปที่ 6-3 Response of a step to a LoG mask.

รูป 6-3(a) แสดงความเข้มของขอบวัตถุของ Ideal step edge ในทิศทาง d_{max} สำหรับ slope ที่วัดตาม zero-crossing ว่าเป็นจุดซึ่งให้ผลของ filter เปลี่ยนค่าผ่าน zero ในบริเวณขนาด 3×3 ค่า magnitude สามารถคำนวณได้ 3 วิธี

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- 1) Local slope คือ ค่า slope ที่มากที่สุด ใน 4 ทิศทางที่เป็นไปได้ ของบริเวณ 3×3
 - 2) The peak to peak difference คือ จุดต่ำสุดและสูงสุดของทั้งสองข้างของ zero-crossing ถูกพบเป็น local slope มากที่สุด [$|p_2 - p_1|$ ในรูปที่ 6-3(b)]
 - 3) Energy คือ ค่าลัมบ์ของพื้นที่ใต้ peak-to-peak curve ในรูปที่ 6-3(b) ก็คือ [$|a_1| + |a_2|$ ในรูปที่ 6-3(b)]
- ซึ่งสองวิธีหลังมีการคำนวณที่ยุ่งยาก ดังนั้นจึงเลือกใช้วิธีแรก

6.3 ตำแหน่งของขอบภาพที่ถูกตัดที่ซีโรครอสซิง

(Edge Position Accuracy at Zero-Crossing)

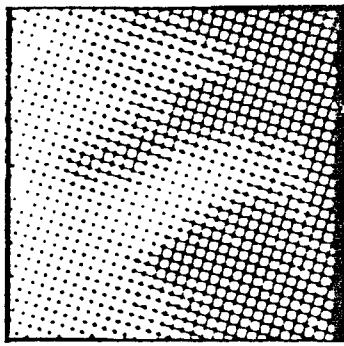
การตรวจจับขอบภาพกับตำแหน่งที่ถูกตัดของขอบภาพขึ้นอยู่กับค่า W เมื่อ W คือ ค่า excitatory region ของตัว operator โดยทั่วไป zero-crossing contour จะมีค่าน้อยกว่าค่า W และอาจอยู่ติดกับ zero-crossing contour ของสภาวะที่ใกล้เคียงกัน อย่างไรก็ตาม ปรากฏการณ์นี้ก็เป็นประโยชน์ในการประยุกต์ใช้งาน ที่มีจุดมุ่งหมายที่จะหาขอบของภาพให้ได้ทั้งหมดรวมทั้งขอบที่มีลักษณะเป็นพื้นผิวด้วย เช่น ในภาพ halftone รูปที่ 6-4

ในรูปที่ 6-4(a) เป็นขอบภาพที่มีลักษณะความสว่างของขนาดเท่ากับ d (solid line) กับภาพของ LoG (broken line) ด้วยค่า central excitatory region เท่ากับ W รูปที่ 6-4(b) แสดงผลตอบสนองของ operator (broken line) แล้วเส้นประในแนวดิ่งเป็นตำแหน่งที่ที่ zero-crossing ถูกตรวจจับและกำหนดให้เป็นขอบ ถ้าค่า $d < W$ ก็จะมีผลตอบสนองที่ $(W-d)/2$ และพบว่าจะมี position energy มากกว่า negative energy (พื้นที่แรเงา)

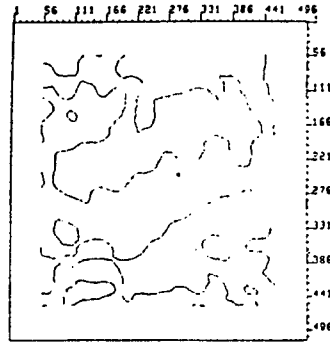
ในรูปที่ 6-5(a) ขนาดของ $d > W$ แต่น้อยกว่า s (s = ขนาดของ operator) รูปที่ 6-5(b) แสดงผลตอบสนองของ operator (broken line) และเส้นประในแนวดิ่ง 2 เส้น แสดงตำแหน่งของขอบภาพ zero-crossing จะใช้แทนตำแหน่งของขอบที่แน่นอนได้ แต่ขอบยังมีขนาดต่อขนาดของขอบอื่นๆด้วย positive energy ยังคงมีค่ามากกว่า negative energy ซึ่งขอบภาพต้องถูกแยกโดยค่า s ที่น้อยที่สุด เพื่อที่จะลดขนาดซึ่งกันและกัน

ขอบที่มีลักษณะเป็นเส้นตรงและหลังคา (ramp and roof) พบได้มาก second derivative operator จะมีผลตอบสนองต่อ "zero" เป็นค่าเดียวภายในเส้นตรงที่มีลักษณะเรียบ ถ้าขนาดของ operator มีค่ามากกว่าหรือเท่ากับความกว้างของเส้นตรง ดังนั้นขอบของเส้นตรงที่กว้างที่สุด ควรจะไม่มากกว่าค่าของ operator (s)

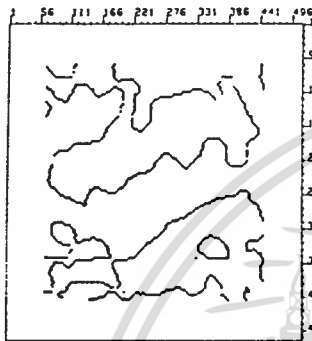
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(a)



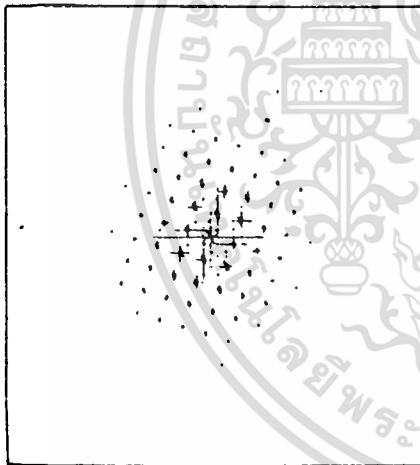
(b)



(c)



(d)



(e)



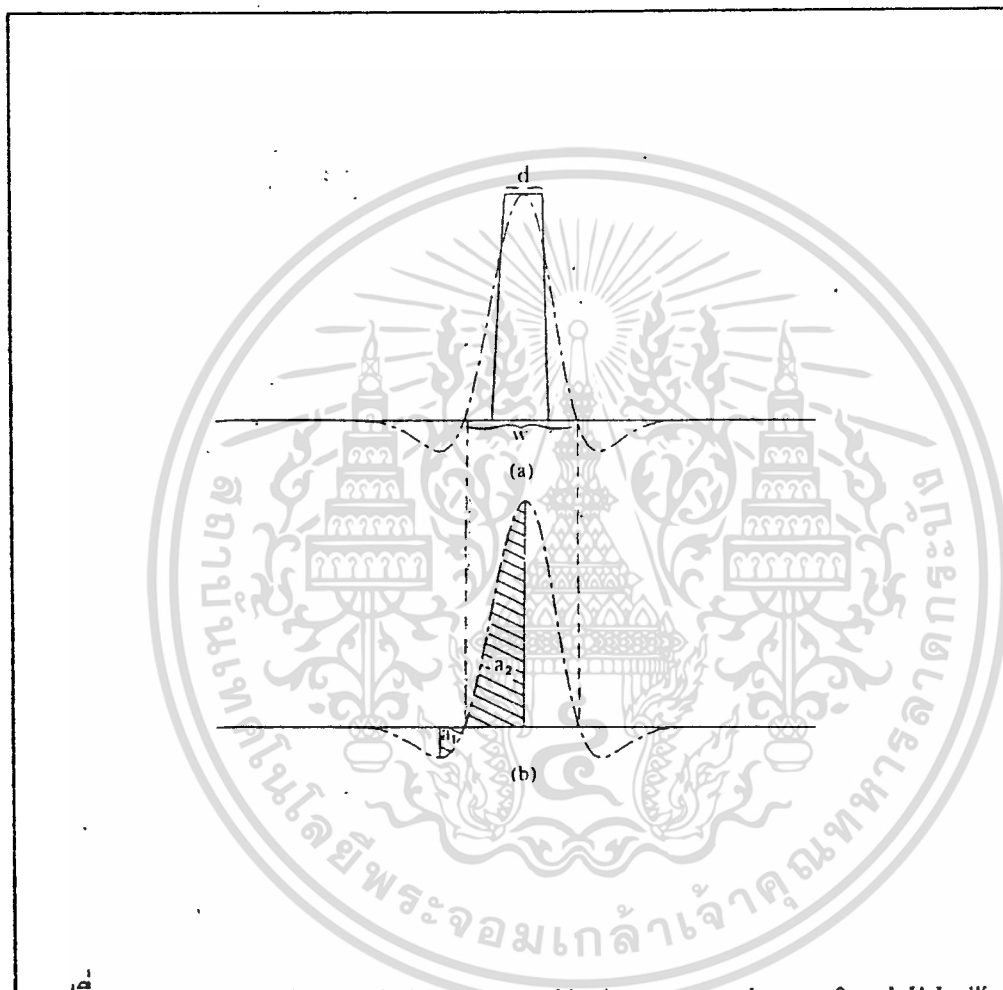
(f)

รูปที่ 6-4 Halftone Image : (a) Gray level image (512 x 512 pixels). (b) Zero-crossings at pixel accuracy from $\sigma = 16.3$. (c) Zero-crossing at pixel accuracy from reduced image (128 x 128 pixels) and a LoG with $\sigma = 4.2$. (d) subpixel precision zero-crossing from (c) with 1/4 th pixel accuracy. (e) The quantity $\log(1+f^2)$, where f is the spectrum , for the full resolution (512 x 512 pixels) halftone image. The quantities were threshold at 1/255 the value of the peak. (f) The same quantity for the reduced version of the halftone image (128 x 128) pixels.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับกรใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์อื่น การค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

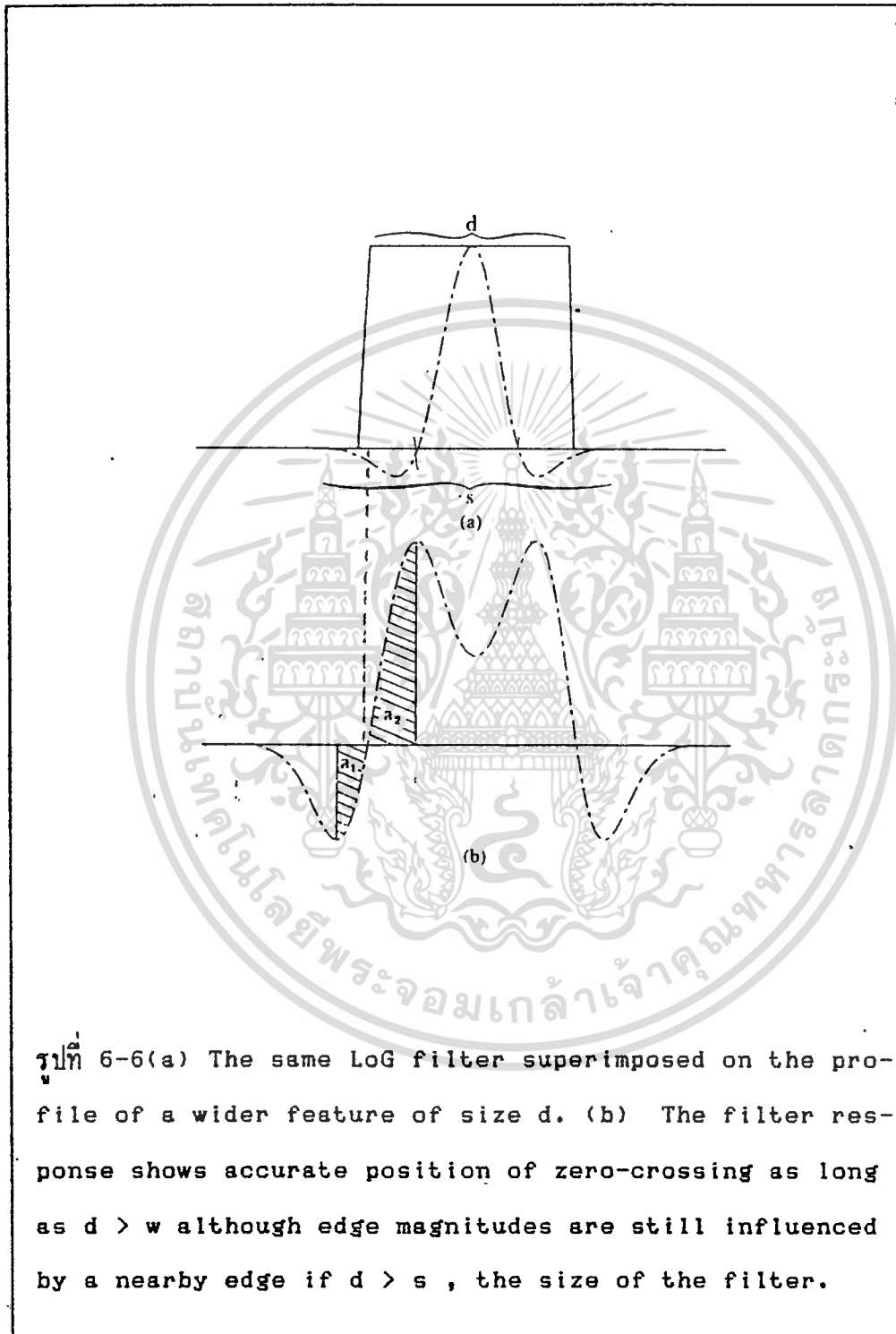
ภาพของ Log รวมทั้ง W และ s ได้แสดงในรูป 6-6(a) พร้อมกับภาพการเปลี่ยนแปลงของ ramp ผลตอบสนองของ operator ได้แสดงไว้ในรูปที่ 6-6(b) (broken line) ขอบของภาพที่ตรวจจับได้แสดงในรูปที่ 6-6(b) โดยเส้นประในแนวตั้งยาวเท่ากับความกว้าง d ของขอบซึ่งมีค่าน้อยกว่าขนาด s มิฉะนั้นขอบจะไม่ถูกกำหนด เมื่อไม่ตรวจจับ zero-crossing ดังเช่นในรูป 6-7(a) และ (b)



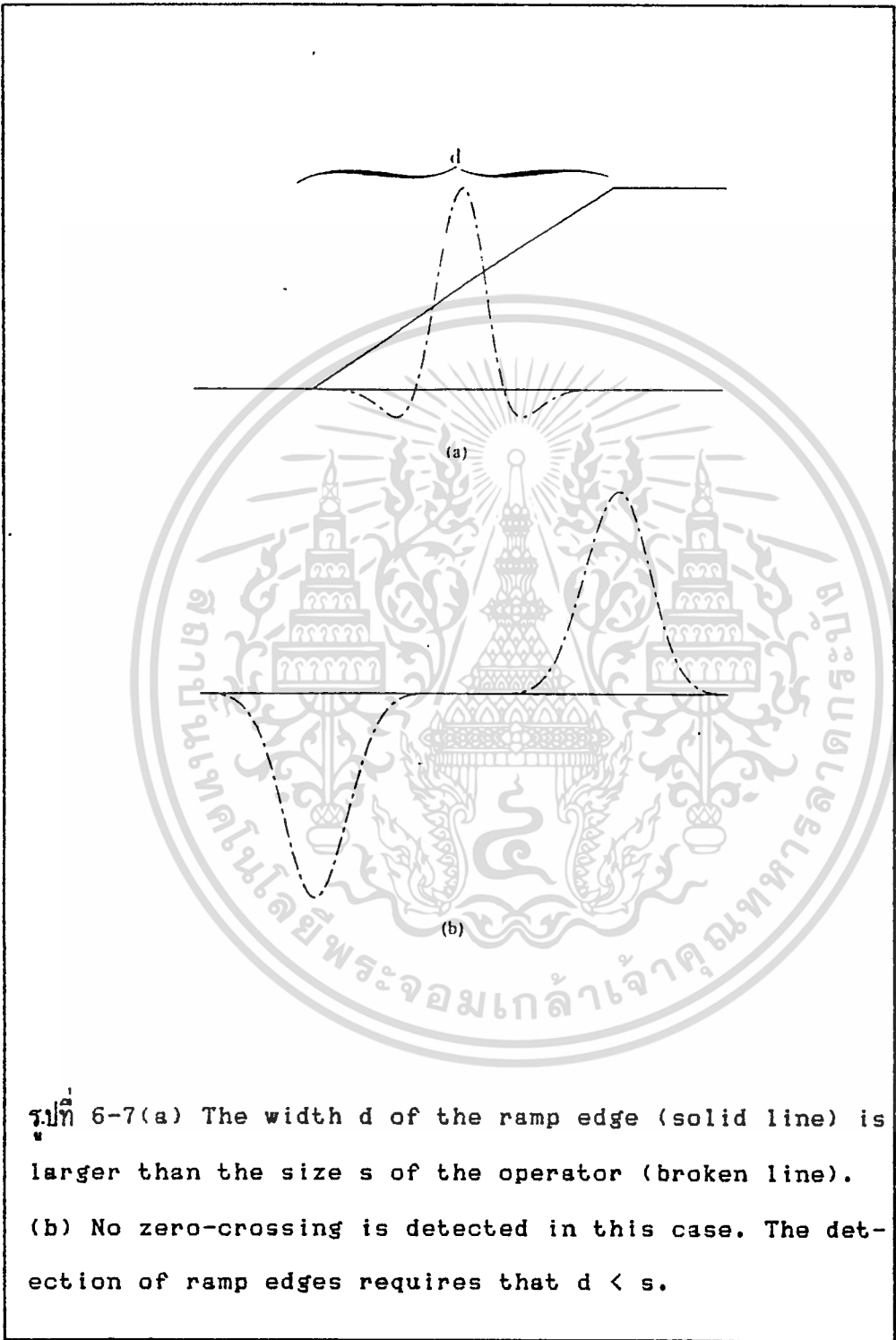
รูปที่ 6-5 (a) A Log with an excitatory region of width W (broken line) superimposed on the profile of a bright feature of size d . (b) the filter response (broken line) shows that when $d < w$. The zero-crossings are displaced by $(w-d)/2$. As indicated by the position of zero-crossings (vertical dotted lines). Nearby edges have a large influence on each other's magnitude; the positive energy is much larger than the negative energy.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับกรณีใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

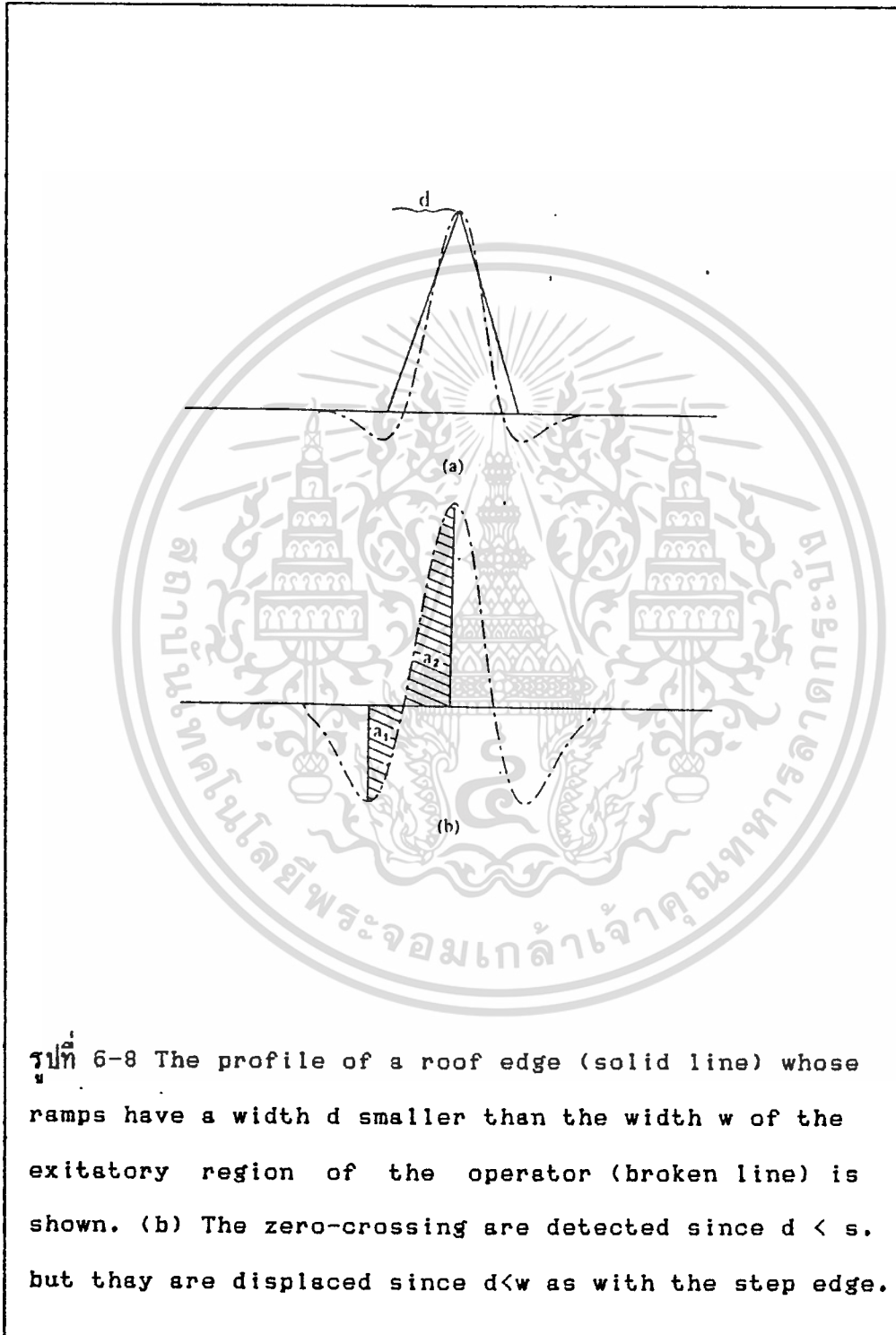


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 6-7(a) The width d of the ramp edge (solid line) is larger than the size s of the operator (broken line).
(b) No zero-crossing is detected in this case. The detection of ramp edges requires that $d < s$.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

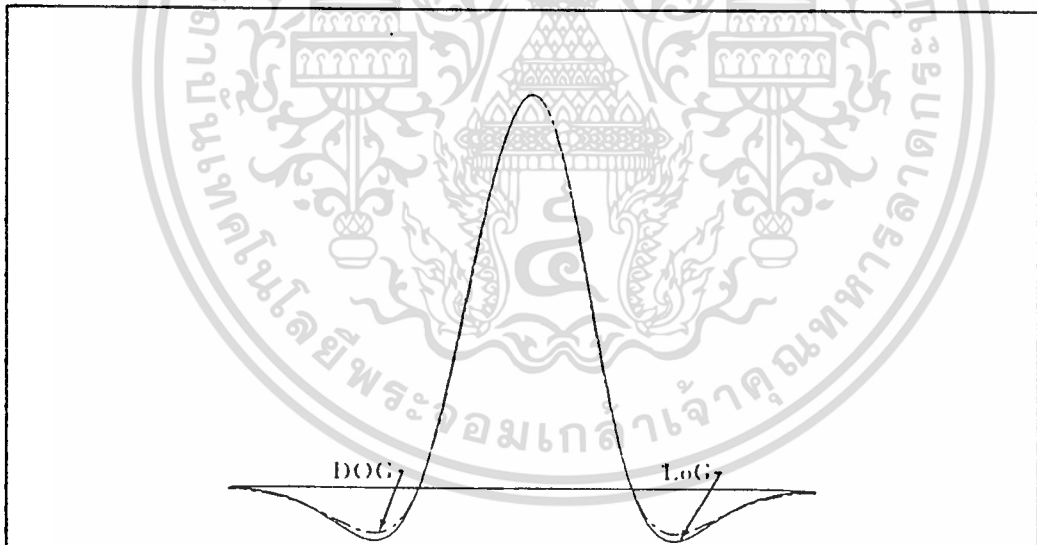


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ผลตอบสนองของ operator ต่อ roof edge ที่แคบ ในรูปที่ 6-8(a) ถูกแสดงในรูปที่ 6-8(b) ถ้าความกว้าง d ของ ramp มีค่าน้อยกว่าค่า w zero-crossing จะตอบสนองโดยการแทนด้วยค่า $(w-d)/2$ เช่นเดียวกับในกรณี step edge ถึงแม้ว่าอิทธิพลของ ramp แต่ละ ramp บนขนาดของขอบแต่ละอัน จะชัดเจนน้อยกว่าก็ตาม

6.4 Separable Decomposition of LoG Operator

โดยปกติจะใช้ operator ที่มีขนาดใหญ่ เพื่อความสะดวกของอัลกอริทึมของการหาขอบ เราจะใช้ operator ขนาดเกินกว่าพิกเซล 5×5 ในทฤษฎีนี้ เพื่อที่จะสนับสนุน operator ขนาดใหญ่ๆ การทำ convolution ด้วย operator ขนาดใหญ่ๆจะทำให้เสียเวลามาก ซึ่งในอดีตเราใช้การประมาณ LoG operator ด้วย Difference of two Gaussian (DoG) ซึ่งจะต่างกันที่ค่า space constant เหตุผลสำคัญที่ใช้ DoG แทน LoG ก็เพราะลักษณะทางกายภาพนั่นเอง ซึ่งดูได้จากรูปที่ 6-9 ที่แสดงให้เห็นถึงความใกล้เคียงกัน



รูปที่ 6-9 A LoG filter (solid line) with $\sigma = 12.35$ and the closest possible DOG (broken line), corresponding to $\sigma = 10$ and $\sigma = 16$, in the optimal ratio proposed in [5]. The DOG requires a slightly larger support ($4W$) than the LoG ($3W$).

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับกรใช้ภายในเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แต่ในปัจจุบันเราใช้การทำงานของคอมพิวเตอร์ โดยแยก Gaussian เป็น 2 ส่วน แต่ส่วนถือเป็น 1 มิติคือ แถวและหลัก ซึ่งจะทำให้เวลาที่ใช้ลดลงจาก M^2 (ในการทำ convolution ปกติ) เป็น $2M$ (โดยการวิธีการ Separate) เราจึงพบว่าการทำงานไม่มีความจำเป็นอีกต่อไป และได้ทำการแยกฟังก์ชัน LoG ออกเป็น 2 ส่วนเป็น row filter และ column filter

$$\nabla^2 G(x, y) = K \left(2 - \frac{x^2 + y^2}{\sigma^2} \right) \exp \left[-\frac{(x^2 + y^2)}{2\sigma^2} \right] \quad (2)$$

where K is a scale factor. Equation (2) can be rewritten as the sum of two separable filters, $h_{12}(x, y)$ and $h_{21}(x, y)$, as follows:

$$\nabla^2 G(x, y) = h_{12}(x, y) + h_{21}(x, y) \quad (3)$$

where

$$h_{12}(x, y) = h_1(x)h_2(y) \quad (4)$$

$$h_{21}(x, y) = h_2(x)h_1(y) \quad (5)$$

$$h_1(\xi) = \sqrt{K} \left(1 - \frac{\xi^2}{\sigma^2} \right) \exp \left[\frac{-\xi^2}{2\sigma^2} \right] \quad (6)$$

$$h_2(\xi) = \sqrt{K} \exp \left[\frac{-\xi^2}{2\sigma^2} \right] \quad (7)$$

ค่า K ที่เลือกใช้จะเป็นสัมประสิทธิ์สูงสุดของ filter ซึ่งสามารถหาได้จากจำนวนบิตของความเที่ยงตรง ผลรวมของสัมประสิทธิ์ของ filter มีค่าเป็น "0" การคำนวณจะลดลงโดยการแยกการทำ convolution ทั้ง 2 เพื่อที่จะลดการเข้าถึงข้อมูล และยังได้เปรียบที่ filter จะมีลักษณะสมมาตรทั้ง row และ column ซึ่งลดจำนวนครั้งของการคูณโดยตัวประกอบอื่นๆให้เหลือ 2

6.5 อัลกอริทึมพื้นฐานการทำนายเพื่อกำหนดตำแหน่งขอบและทิศทางของภาพ

(Predicate-Based Algorithm For Determining Edge Location and Orientation)

เมื่อใช้ Second Derivative Edge Operator ดังเช่น LoG ขอบจะถูกจับได้ โดยตำแหน่งที่ค่าของเอาต์พุตของการ Convolution เปลี่ยนเครื่องหมายเป็นตรงกันข้าม (โดยที่ Zero-Crossing จะอยู่ตรงกลาง) สำหรับการให้ประโยชน์ อย่างเช่น ระบบสเตอริโอที่ใช้ Epipolar Geometry Zero-crossing อาจจะถูกจับได้โดยการสังเกตภาพที่ได้จากการคำนวณเอกรายนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แนวแล้วทางแนวนอน สำหรับสองพิกเซลที่ติดกันที่เครื่องหมายตรงกันข้าม หรือ 3 พิกเซลที่ติดกันที่ตรงกลางเป็น "0" และอีกสองพิกเซลมีเครื่องหมายตรงกันข้าม อย่างไรก็ตาม เมื่อใช้การตรวจจับ 2 มิติ ต้องสนใจการทำงาน เพื่อให้แน่ใจว่าเส้น Zero-crossing รักษาขนาดและรูปร่างของระยะเครื่องหมายตรงกันข้ามอันที่กันขอบเขตไว้ และป้องกันการผิดพลาด เช่น การสร้างขอบที่ไม่ต้องการตามแนวทแยง การพิจารณาส่วนนี้จึงเป็นสิ่งสำคัญ เมื่อใช้ operator ขนาดเล็ก และบริเวณภาพที่คำนวณเพิ่มขึ้น

A.Huertas และ D.King ได้พัฒนาประสิทธิภาพ Algorithm พื้นฐานการทำงาน เพื่อกำหนดจุดแกนในสองทิศทาง โดยหลักการนี้ แต่ละย่าน 3×3 ในภาพที่คำนวณถูกตรวจว่า เข้าหลักเกณฑ์แบบใดใน 11 แบบหรือไม่ อันที่กำหนดตำแหน่งของขอบได้ 24 ตำแหน่งการทำงานแบบนั้นแสดงในบทส่งท้าย เครื่องหมายบวก และ/หรือ ลบ ใน Template แสดงว่า ค่าบวกและ/หรือ ค่าลบ ต้องแสดงที่ตำแหน่งนั้น ในภาพที่ผ่านการ filter แล้วอักษรแสดงที่ตำแหน่งแสดงค่าที่ถูกทดสอบ ตำแหน่ง ของขอบแสดงในหน้าต่าง แสดงโดยพิกเซลวงกลม และสัมพันธ์กับ ทิศทางที่ถูกแสดงโดยลูกศรที่อยู่ถัดไป

ถ้าหน้าต่างขนาด 3×3 ในภาพที่ผ่านการกรอง (filter) แล้วเข้าได้กับการทำนายแบบใดแบบหนึ่งขอบจะถูกมาร์ค (mark) ที่พิกเซลที่เหมาะสม และขนาดและทิศทางของขอบที่สัมพันธ์จะถูกคำนวณ การทำนายทั้ง 11 แบบ จะถูกสร้างมาโดยเฉพาะเพื่อที่จะรักษา topology (รูปแบบในการติดต่อสื่อสารข้อมูล) ของบริเวณเครื่องหมาย และป้องกันการสร้างขอบที่ไม่พึงประสงค์ เมื่อ Template นี้มีรูปแบบที่ชัดเจน เครื่องหมายของค่าพิกเซลในหน้าต่าง 3×3 ยังเป็นเงื่อนไขบังคับทิศทางของขอบ (tangent ของบริเวณ Zero-crossing) ในวิธีการนี้เราจะหาทิศทางโดยทั่วไปที่มี 8 ทิศทาง และใช้การหาทิศทางโดยการทำนาย เพื่อที่จะจำกัดทางเลือกของทิศทาง 3 ใน 8 ทิศที่เป็นไปได้ ซึ่งเป็นทิศทางหลักของขอบที่ทำนายได้ (ดูได้ในบทส่งท้าย) และอีกสองทิศทางที่กำมุก 45 องศาทางด้านอื่นๆ ทิศทางสุดท้ายที่หาได้จะเป็นหนึ่งทิศทางซึ่งมีความแปรปรวนที่น้อยที่สุดในทิศทางของมุมของมัน เราพบว่าเทคนิคนี้ ทำให้การประมาณทิศทางของขอบภาพ มีค่าดีกว่าการคำนวณทิศทางอย่างง่ายรอบย่าน 3×3 ของมันที่มีการกำหนดทิศทางของค่าทิศทางที่มากที่สุด และการเลือกทิศทางของมัน สำหรับรายละเอียดเพิ่มเติมของอัลกอริทึมของการตรวจจับ Zero-crossing อยู่ในบทส่งท้าย

การคำนวณหา Zero-crossing ด้วย Laplacian of Gaussian ต้องทำตามขั้นตอนดังนี้

1. convole ภาพ I ขนาด $R \times C$ ด้วย LoG filter ขนาด M^2 แล้วจะได้ภาพที่ถูก filter เป็น I_{LoG} ขนาด $R \times C$

เอกสารนี้เป็นเอกสารที่เผยแพร่โดยทางโรงเรียนเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. หาภาพ Zero-crossing ให้ชื่อเป็น $I_{\text{Zero-cross}}$ ขนาด $R \times C$ โดยการหาตำแหน่งของ Zero-crossing ใน I_{Log} โดยใช้อัลกอริทึมที่ได้กล่าวไว้แล้ว
3. จากค่าของ Zero-crossing แต่ละตัวและย่านรอบๆ อีก 8 ตำแหน่ง นำมาคำนวณหาทิศทาง โดยใช้การทำนายขอบที่แสดงไว้ใน

6.6 สรุป

จากที่กล่าวมาทั้งหมดสามารถสรุปได้เป็นดังนี้

- มีหลายวิธีที่จะหาขอบภาพ เช่น การทำ Interpolation ของ subpixel precision และ facet model ที่ดูเหมือนว่าจะเหมาะสม
- การใช้ subpixel precision บางครั้งอาจจะได้ภาพที่ไม่ถูกต้อง ซึ่งพบได้จากขอบของภาพที่มีการเชื่อมเกี่ยวกันระหว่าง procedure
- การตรวจหา global หรือ macro edge ในนามจะต้องใช้ convolution mask ที่มีขนาดใหญ่และใช้เวลามาก การที่จะประหยัดเวลาทำได้โดย แยก LoG mask ออกเป็นสองส่วนดังที่กล่าวมาแล้ว

6.7 บทส่งท้าย

(A) : Example of Separable Decomposition of a LoG Filter

ตามตัวอย่างนี้ จะอธิบายการแยกส่วนของ LoG Filter ขนาดเล็กๆ 11×11 ที่มีค่า $\sigma = 1$ และค่า $K = 256$ เป็น row filter และ column filter ขนาด 1 มิติ แล้วนำมาสร้างใหม่เป็น 2-D LoG Filter ในที่นี้จะแสดงเฉพาะด้านบนซ้ายของ filter ด้วยขนาด 1 ใน 4 เท่านั้น เพราะมีลักษณะเป็น Rotational symmetric เหตุผลที่ใช้ค่า K มีค่าเท่ากับ 256 นั้นก็เพราะว่า ค่าสัมประสิทธิ์ของ peak filter ถูกแทนด้วย 8 bit precision และผลรวมของสัมประสิทธิ์ของ filter เป็น "0" ค่า K เป็นค่า scale constant ไม่มีความสัมพันธ์โดยตรงกับขนาดของ operator แต่อย่างไรก็ตาม ค่า scale constant ค่าใหญ่ มักจะถูกใช้ และจะไม่มีการตัดขอบอย่างไม่มีเหตุผลของสัมประสิทธิ์ก็จะไม่เกิดขึ้นและทำให้ยากต่อการสร้าง operator ที่สมดลกัน

สาเหตุที่เกิดความแตกต่างที่น้อยมากๆ ระหว่างค่าสัมประสิทธิ์ของภาพที่ได้จากการสร้างขึ้นใหม่ (Reconstruction) ด้กับภาพที่ได้จากการใช้ filter ดั้งเดิม (Original filter) ค่า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เป็นผลอันเนื่องมาจาก quantization effect ของการแทนที่สัมประสิทธิ์ของ row และ column filter ที่เป็นจำนวนจริงนั่นเอง

Example : Row and Column Filter

$$h_2() = [0 \ 0 \ 0 \ -6 \ 0 \ 16 \ 0 \ -6 \ 0 \ 0 \ 0]$$

$$h_1() = [0 \ 0 \ 0 \ 2 \ 10 \ 16 \ 10 \ 2 \ 0 \ 0 \ 0]$$

จะได้ Reconstruction Laplacian-Gaussian Filter จากการให้ Separate LoG Filter โดยใช้สมการที่ 2,3,4,5,6 และ 7 เป็นต้น (แสดงเฉพาะ upper left quadrant)

0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	-24	-60	0
0	0	0	-60	0	160
0	0	0	0	160	512

(B) : Predicate of edges

ในอัลกอริทึมนี้ใช้หา Zero-crossing ใน 2 ทิศทางขนาด 3 x 3 ในภาพที่คำนวณแล้ว ถูกหาว่าเข้ากับแบบใดในทั้งหมด 11 แบบของการทำนาย Zero-crossing ภายใต้แบบต่างๆ จาก A-K ที่กำหนดตำแหน่งขอบได้ 24 ตำแหน่ง (การรวมกันอย่าง nonredundant ของเครื่องหมายพิกเซลในช่วง 3 x 3 และ ตำแหน่ง Zero-crossing) เครื่องหมายบวก และ/หรือ ลบใน template แสดงว่าค่าบวก และ/หรือ ค่าลบต้องแสดงที่ตำแหน่งในภาพที่ผ่านการกรองแล้วเพื่อจับคู่กัน ตำแหน่งหน้าตาต่างที่ให้เป็น a-i จากซ้ายไปขวา บนไปล่าง ตัวอักษรแสดงว่าตำแหน่งที่ให้มาแสดงค่าที่ถูกทดสอบ ตำแหน่ง ของขอบ ในหน้าตาต่างถูกแสดงโดยพิกเซลที่มีวงกลมและสัมพันธ์กับ ทิศทาง ที่แสดงโดยลูกศรที่อยู่ถัดไปเท่านั้น ไม่นอนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทำนายประกอบด้วย 2 กลุ่มหลัก กลุ่มแรก (ทำนายแบบ A-D) ทำกับค่าศูนย์ในเอาท์พุทของการ convolution การทำนายแต่ละแบบมีทิศทางดังนี้ ทำนายแบบ A ทิศใต้ แบบ B ทิศเหนือ แบบ C ทิศตะวันตก แบบ D ทิศตะวันออก ส่วนทิศทางที่ทะแยงต้องพิจารณาว่า เป็นแบบแนวตั้งหรือว่าแนวนอน การทำนาย 4 แบบนี้ ตั้งแต่ A-D และครอบคลุมทั้ง 8 ทิศทางที่เป็นไปได้ของ zero-crossing ที่พิกเซลตรงกลาง ตอนแรกเราตรวจสอบเครื่องหมายก่อน เพื่อหาแบบทำนายที่เข้ากัน ถ้าไม่มีแบบใดที่เข้ากันได้แล้ว ก็จะไม่มีการ zero-crossing ที่เหมาะสมกัน แต่ถ้า zero-crossing ถูกจับได้ที่พิกเซลตรงกลาง ดังแสดงโดยพิกเซลที่มีวงกลมในตำแหน่ง e กรณีย่อยของการทำนายที่เข้ากัน อันที่นำไปสู่การเพิ่มการตรวจสอบเครื่องหมายเท่ากับค่าของพิกเซล นำไปใช้เพื่อกำหนดทิศทางขอ

ลองพิจารณาสักหนึ่งตัวอย่าง : ถ้าค่าพิกเซลที่อยู่ทางซ้ายของ "zero" เป็นบวก และค่าพิกเซลที่อยู่ทางขวาของ "zero" เป็นลบ โดยไม่สนใจค่าของพิกเซลตัวอื่นใน window ขนาด 3×3 ว่าจะมีค่าเป็นเช่นไร ดังนั้นคุณสมบัตินี้จะสอดคล้องกับ การทำนายแบบ A (Predicate A) และกรณีย่อย (subpixel) ของ predicate A ถูกใช้ในการหาทิศทาง (direction) ของขอบภาพ จะได้ทิศทางไปทางทิศใต้ แต่ทิศทางของภาพที่ถืออาจหาได้ เมื่อค่าพิกเซล b และ h ถูกนำมาคำนวณด้วย

กลุ่มที่ 2 ของการทำนายขอบ (E-K) คิดที่ค่าบวกเปลี่ยนไปเป็นค่าลบในเอาท์พุทของ convolution เราเลือกกำหนดช่วง 3×3 ของพิกเซลที่มีค่าบวก เพื่อหาการเปลี่ยนไปเป็นค่าลบ การทำนายที่เข้ากันของแบบ E-K เราจะดูที่เครื่องหมายก่อนเหมือนดังกลุ่มที่ 1 โดยที่ไม่จำเป็นต้องไปพิจารณาค่าที่มุมทั้ง 4 ของช่วง 3×3

-ลองพิจารณาการทำนายแบบ E (Predicate E) เป็นตัวอย่าง
กรณี 1 ถ้าพิกเซล e และ f เป็นบวก (positive) และพิกเซล d เป็นลบ (negative)
ดังนั้น

- ถ้าพิกเซล $|h| < |e|$ เมื่อ zero-crossing ถูกพบในตำแหน่ง h และมีทิศทางไปทางทิศตะวันตกเฉียงเหนือ ลำดับต่อไป ถ้าพิกเซล $|d| < |e|$ เมื่อ zero-crossing ถูกพบในตำแหน่ง d และมีทิศทางไปทางตะวันตกเฉียงเหนือ ออกจากการหา (exit)
- ถ้าต่างไปจากนี้ zero-crossing จะถูกพบที่ตำแหน่ง e มีทิศทางตะวันออกเฉียงเหนือ

กรณี 2 ถ้าพิกเซล b เป็นลบและพิกเซล h เป็นบวก ดังนั้น

- ถ้าพิกเซล $|b| < |h|$ เมื่อ zero-crossing ถูกพบในตำแหน่ง b
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

และ มีทิศทางไปทางตะวันออกเฉียงเหนือ ต่อจากนั้น ถ้า $|d| <$ พิกเซล e zero-crossing จะถูกพบที่ตำแหน่ง d และ มีทิศทางไปทางทิศตะวันออกเฉียงเหนือ ออกจากการหา (exit)

- ถ้าต่างไปจากนี้ zero-crossing ถูกพบในตำแหน่ง e และ มีทิศทางไปทางตะวันออกเฉียงเหนือ

กรณี 3 ถ้าพิกเซล b และ h เป็นบวก ดังนั้น

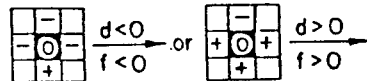
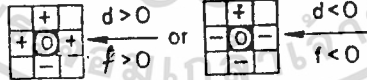
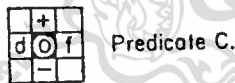
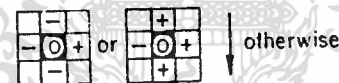
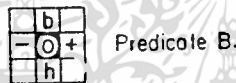
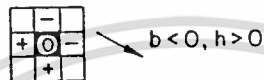
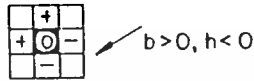
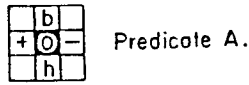
- ถ้าพิกเซล $|d| <$ พิกเซล e ดังนั้น zero-crossing ถูกพบที่ตำแหน่ง d และ มีทิศทางไปทางเหนือ ออกจากการหา (exit)

กรณี 4 ถ้าไม่เข้ากรณีข้างบนทั้งหมด จะ mark ที่ e เป็น zero-crossing กับทิศทางไปทางเหนือ

- ถ้าต่างไปจากนี้ ให้หา Predicate ถัดไป



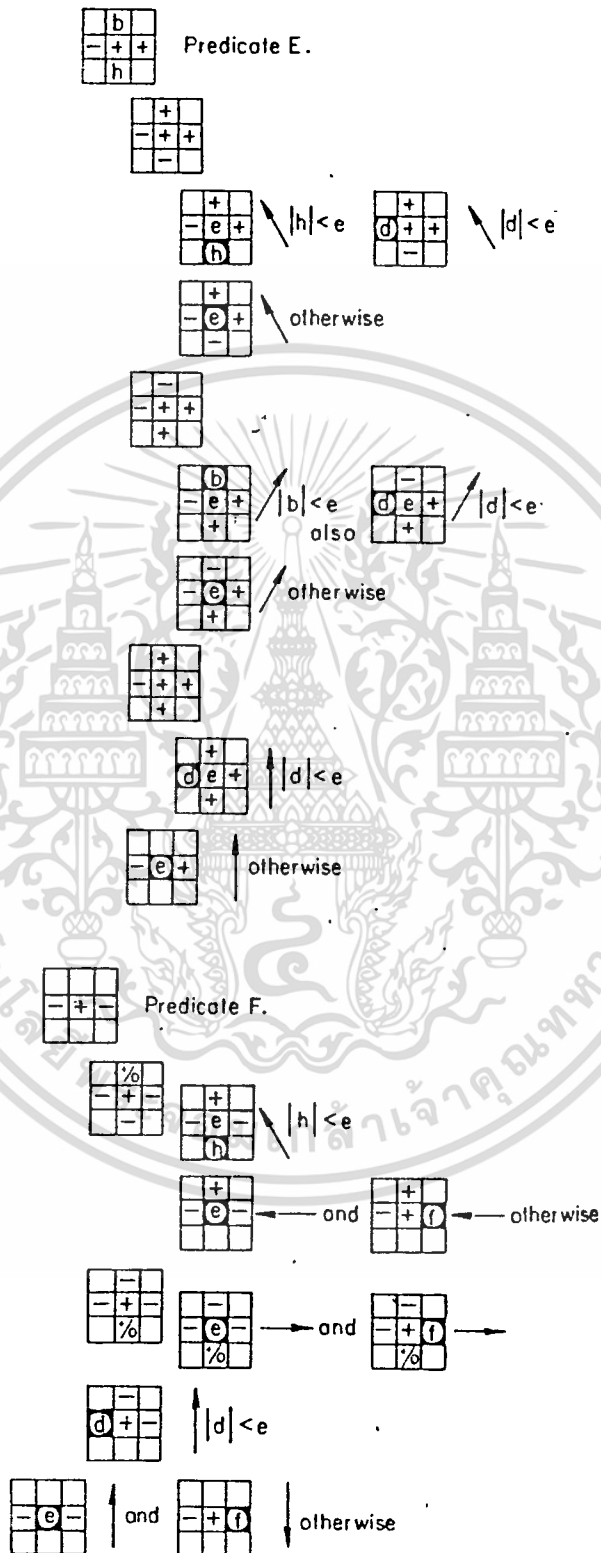
PREDICATE A - D



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์อื่นใด

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

PREDICATE E - F



เอกสารนี้เป็นเอกสารสงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้เผยแพร่ไปใช้ประโยชน์ทางการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

PREDICATE G - J



Predicate G.



$|b| < e$ $\left\{ \begin{array}{l} \rightarrow f > 0 \\ \searrow f < 0 \end{array} \right.$



$f < 0, |f| < e$



$f < 0, |f| > e$



otherwise



Predicate H.



$f < 0$ $\left\{ \begin{array}{l} \nearrow g = 0 \\ \searrow g > 0 \end{array} \right.$



and



otherwise



otherwise



Predicate I.



$|f| < e$



otherwise



Predicate J.



$|h| < e$



otherwise



$c > 0, |f| < e$



otherwise



$c < 0, |f| < e$

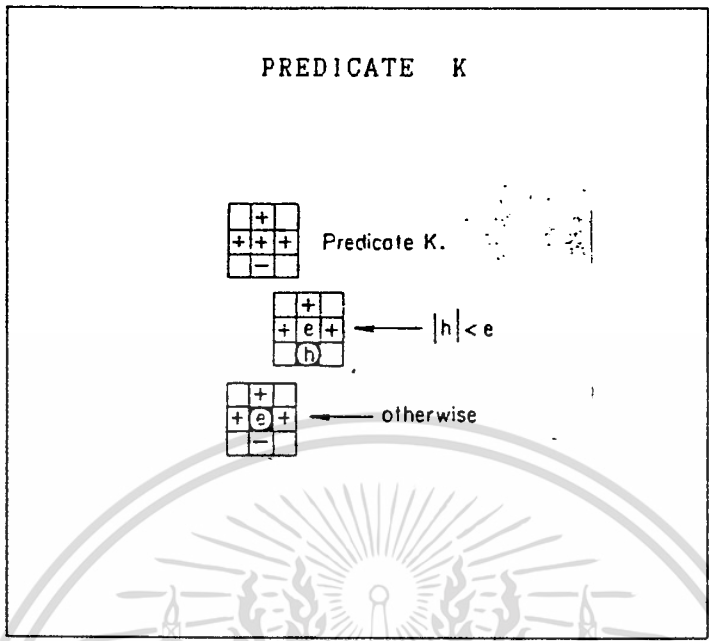


otherwise



$g > 0$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 7

การหาระยะทาง

(Finding Distance)

7.1 Range Finding

Passive Triangulation

Passive Triangulation เป็นการมองเห็นภาพคู่ เราอาจกล่าวได้ว่า การมองเห็นภาพคู่นี้เป็นไปตามหลักชีววิทยา เพราะวิธีนี้เป็นวิธีที่มนุษย์ และสัตว์ส่วนมากมองเห็นที่มีความลึกทำนองเดียวกัน อุปกรณ์ที่ให้ภาพ 2 ภาพ จะทำให้เรารู้อส่วนของระยะทางจากวัตถุถึงกล้อง ดังแสดงในรูปที่ 7-1 ซึ่งเป็นระบบการมองเห็น อุปกรณ์การให้ภาพ 2 ภาพนี้อาจจะเป็น TV 2 เครื่อง diode matrix หรือ กล้อง CCD เพราะว่าตามหลักชีววิทยา การมองเห็นภาพโดยสมบูรณ์ต้องมีสองตา

การทำงานของระบบตามรูปที่ 7-1 เราจะต้องรู้ระยะห่างของกล้องสองตัว (d) และระยะโฟกัสของกล้อง (f) ส่วนที่ต้องคำนวณคือ ระยะจากกล้องถึงวัตถุ (r_1 and r_2) กล้องทั้งสองตัวก็จะให้จุดต่างๆ ของภาพ เช่น จุด P จะมีสอง pixel ที่แทนจุดนี้ pixel แรกจะอยู่ในภาพทางด้านซ้าย และอีก pixel หนึ่งจะอยู่ในภาพทางด้านขวา

ถ้าภาพของกล้องทั้งสองเหลื่อมกัน จุดของภาพทั้งสอง (x_1 and x_2) จะไม่ตรงกัน ดังนั้นเราสามารถหาระยะทางจากกล้องถึงวัตถุได้จากความแตกต่างกันตรงนี้ ค่าความแตกต่างนี้ คำนวณได้เป็นค่า Absolute ผลของความแตกต่างเราเรียกว่า "Disparity" ระหว่างจุดเดียวกันของภาพทั้งสอง

ระยะจากกล้องถึงวัตถุ (r) จะเป็นส่วนกลับของความแตกต่างระหว่าง x_1 กับ x_2 เช่นค่าของความแตกต่างระหว่าง x_1 กับ x_2 เป็น "0" (x_1 และ x_2 เป็นจุดเดียวกัน) ค่าระยะจากกล้องถึงวัตถุจะเป็น infinite ดังนั้น ระยะทางระหว่างจุดของภาพทั้งสองเป็น "0" ในทางกลับกัน ระยะจากกล้องถึงวัตถุ มีค่าน้อยๆ ซึ่งจะได้ความแตกต่างของ x_1 กับ x_2 มีค่ามาก

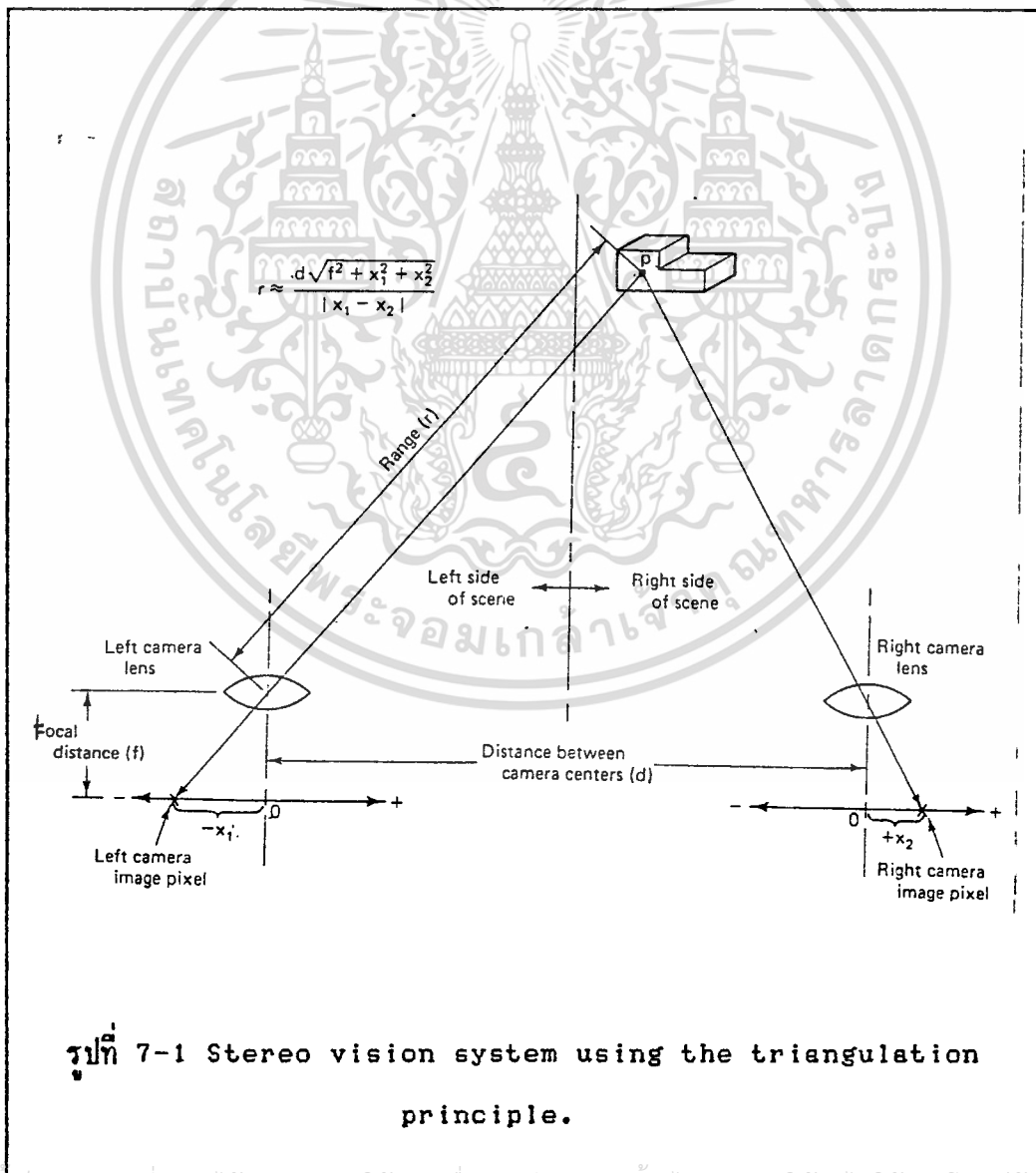
ระบบการมองเห็นภาพคู่ ในรูปที่ 7-1 เป็นการหาระยะของกล้องถึงวัตถุ สามารถคำนวณได้ตามสมการข้างล่างนี้

$$\text{Range} = r = \frac{d \sqrt{f^2 + x_1^2 + x_2^2}}{|x_1 + x_2|}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้ไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- โดยกำหนดให้ r = ระยะจากกล้องด้านซ้ายถึงวัตถุ หรือ ระยะจากกล้องด้านขวาถึงวัตถุ
 d = ระยะห่างของกล้องทั้งสอง
 f = เป็นความยาวโฟกัสของกล้อง
 x_1 = ระยะทางของจุดของภาพที่ตกกระทบ image ด้านซ้าย
 x_2 = ระยะทางของจุดของภาพที่ตกกระทบ image ด้านขวา

จากสมการที่ใช้ในการหาระยะ r สามารถเป็นระยะทางจากวัตถุถึงกล้องด้านซ้าย หรือ ระยะจากวัตถุถึงกล้องด้านขวา ขึ้นอยู่กับว่าวัตถุนั้นอยู่ในตำแหน่งด้านใดของกล้อง ตัวอย่างเช่น ถ้าวัตถุอยู่ทางด้านซ้ายของกล้อง r จะเป็นระยะจากกล้องถึงกล้องทางด้านซ้าย



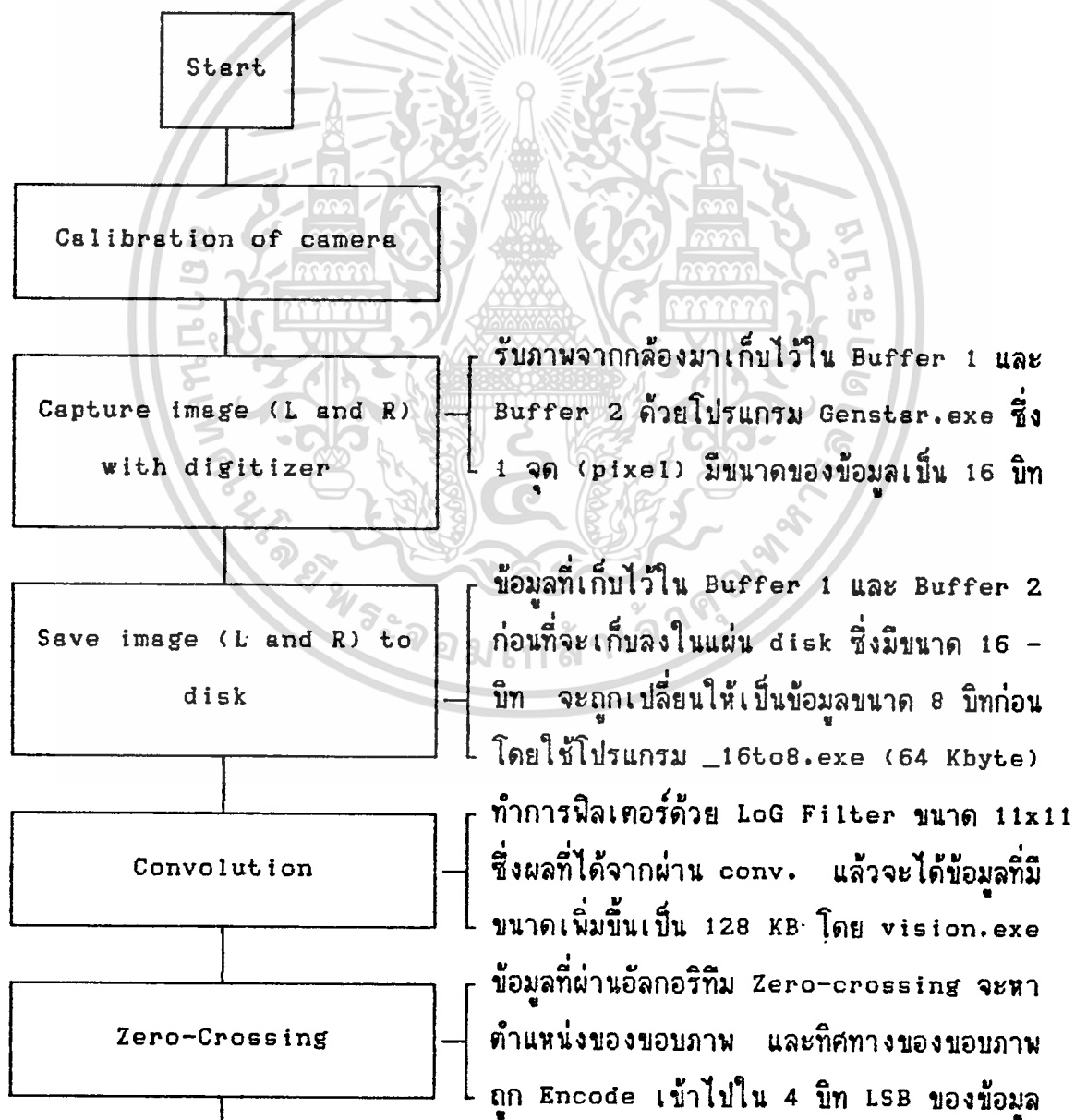
บทที่ 8

สเตอริโอวิชั่นและการทดลอง

(Stereo Vision and Experiment)

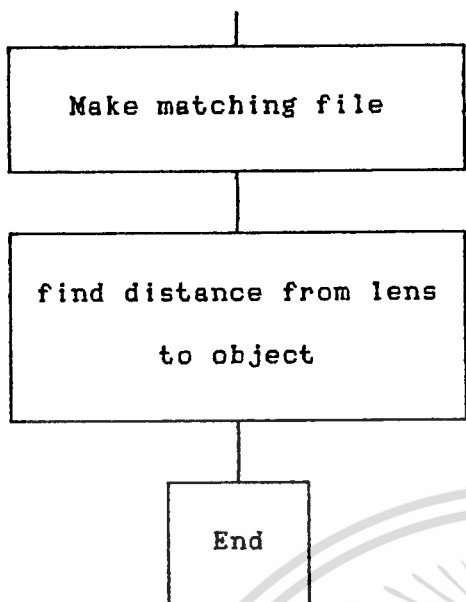
8.1 Stereo vision and experiment

ในบทนี้เราจะนำเอาทฤษฎีต่างๆ ที่ได้กล่าวมาแล้วในบทที่ 1 ถึง บทที่ 7 มาใช้งาน ซึ่งหลักการของสเตอริโอวิชั่นนั้น เราขออ้างถึงรูปที่ 7-1 ในบทที่ 7 และสูตรในการคำนวณ ในบทที่ 7 หน้า 7-1 จะพบว่าพารามิเตอร์ที่เรายังไม่ทราบค่าโดยตรงอยู่ 2 ตัว คือ x_1 และ x_2 ซึ่งเราจะได้อธิบายในภายหลัง ในการทำงานของสเตอริโอวิชั่นนั้น เราสามารถเขียนขั้นตอนการทำงานเป็น flow chart ได้ดังนี้



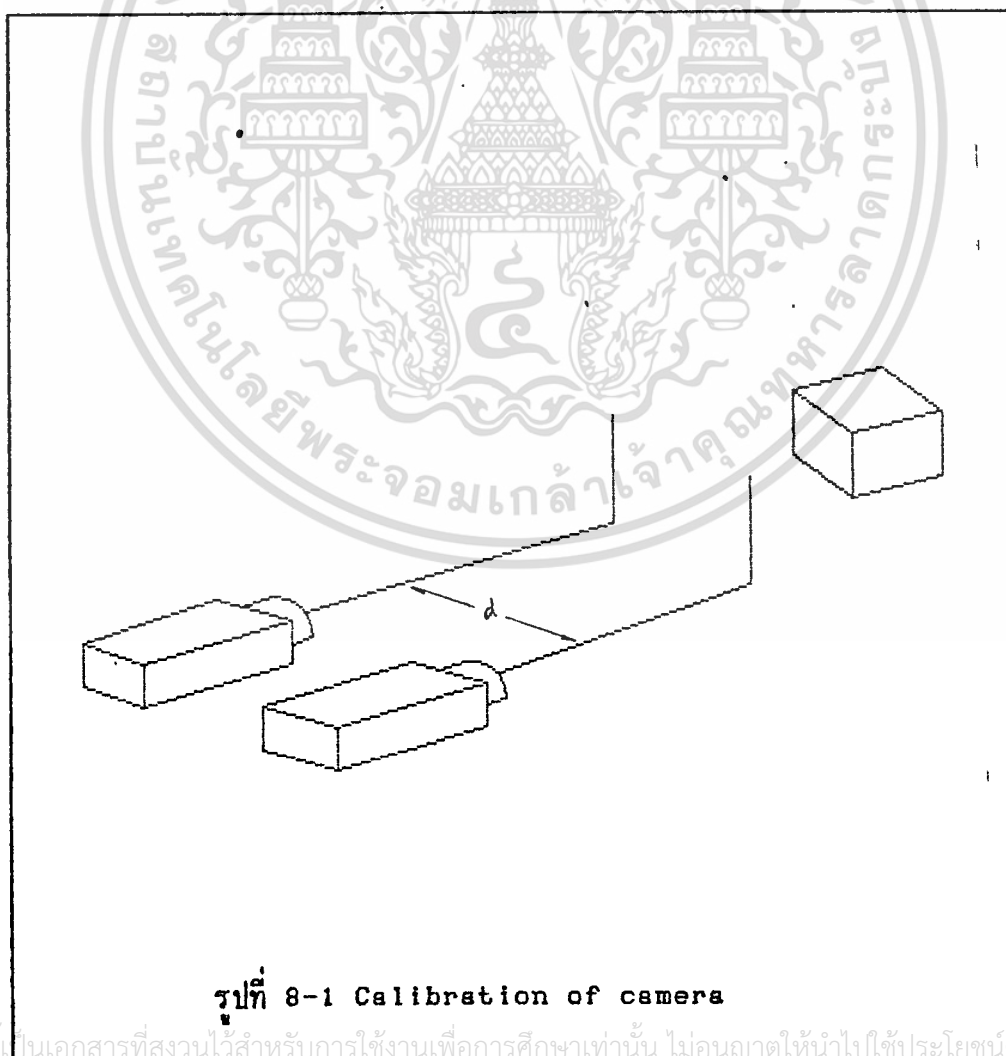
ต่อหน้าถัดไป

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ข้อมูลที่ได้จากขั้นตอนนี้จะได้ file ขนาด 8 K
อีก 1 ไฟล์

Stereo Vision Flow Chart



รูปที่ 8-1 Calibration of camera

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้ทำไปใช้ประโยชน์ทางการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Calibration of camera

ในการถ่ายภาพเก็บเอาไว้ในแต่ละครั้งนั้น เราจำเป็นต้อง calibrate กล้องเสียก่อนทุกครั้ง เนื่องจากการทดลองเราใช้กล้องเพียงตัวเดียว ดังนั้นการถ่ายภาพ ซ้ายและขวา เราจะต้องทำให้ภาพข้างซ้ายและขวายู่ขนานกันจริงๆ ดังแสดงในรูป 8-1 และดำเนินการตามขั้นตอนต่อไปนี้

1. ทำการตีเส้นขนานกัน 2 เส้น ซึ่ง 2 เส้นนี้จะต้องขนานกันและมีระยะห่างเท่ากับ D
2. จากนั้นพยายามจับภาพเส้นตรงทางซ้ายให้อยู่กึ่งกลางของจอภาพ
3. ทำการ capture ภาพมาเก็บไว้ใน Image card ที่ buffer 1
4. จากนั้นพยายามจับภาพเส้นตรงทางขวาให้อยู่กึ่งกลางของจอภาพ
5. ทำการ capture ภาพมาเก็บไว้ใน Image card ที่ buffer 2

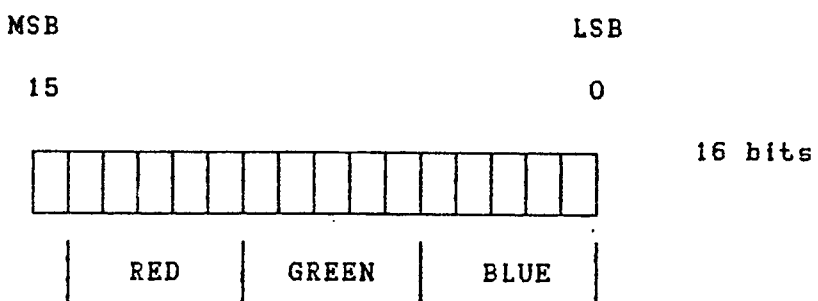
ในการทดลองเราใช้ระยะโฟกัสของเลนส์ $f = 7 \text{ mm}$. (เราปรับโฟกัสไปที่ 7 mm) ส่วนของ D เราเลือกใช้ $10,8,6 \text{ cm}$.

Capture image with digitizer

เป็นขั้นตอนที่นำข้อมูลที่ได้จากกล้องมาเก็บไว้ในหน่วยความจำ (memory) ของคอมพิวเตอร์ ซึ่งในที่นี้เราใช้ image card ของ Eyegrabber กับโปรแกรม genstar.exe ทำหน้าที่ในการ Capture ภาพ

Eyegrabber card

Eyegrabber card จะมี buffer สำหรับเก็บภาพอยู่ 2 buffer ในที่นี้เราเราจะให้ buffer 1 ใช้เก็บภาพซ้าย (left image) และให้ buffer 2 ใช้เก็บภาพขวา (right - image) โดยที่ buffer 1 จะแสดงภาพได้ขนาด 512×256 จุด ใน 1 จุดจะมีข้อมูลเป็น 16 บิต ซึ่งใน buffer จะมีรูปแบบการเก็บข้อมูลเป็นดังนี้

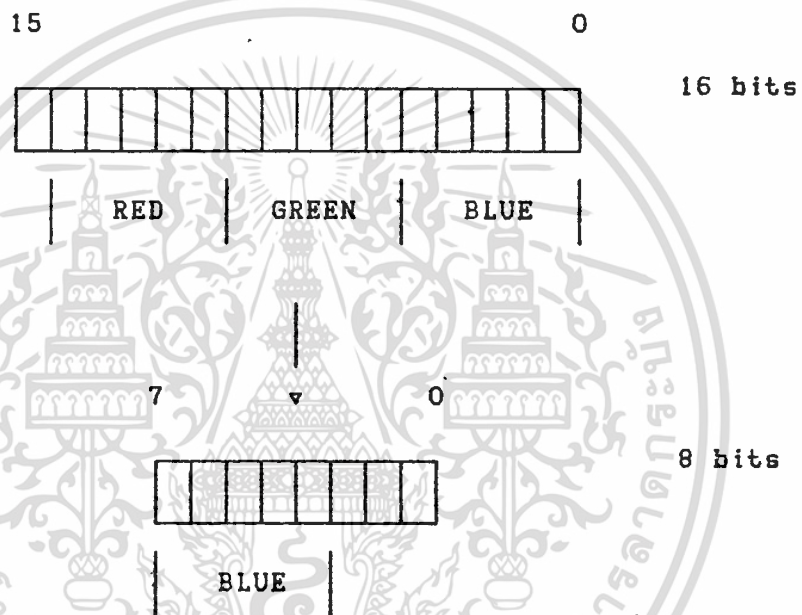


โดยจะมีโปรแกรม genstar.exe เป็นตัวทำหน้าที่นำภาพที่ได้จากกล้องมาเก็บไว้ใน buffer 1 และ buffer 2

หมายเหตุ ในการถ่ายภาพแต่ละครั้งควรจะปรับภาพให้เป็นภาพขาว-ดำเสียก่อน จึงเก็บภาพลง buffer

Save image to disk

เป็นขั้นตอนที่รับภาพจาก buffer 1 (L) และ buffer 2 (R) เอาไปเก็บไว้ในแผ่น disk โดยใช้โปรแกรม `_16to8.exe` ทำหน้าที่นี้ `_16to8.exe` นอกจากจะทำหน้าที่รับภาพจาก buffer 1,2 ไปเก็บไว้ใน disk แล้ว ยังมีหน้าที่อีกอย่างหนึ่งคือ เปลี่ยนข้อมูลจาก 16 บิตไปเป็น 8 บิต



โปรแกรม `_16to8.exe` จะทำการอ่านข้อมูลภาพจาก แอดเดรสที่ `d000:0000` ซึ่งเป็นแอดเดรสเริ่มต้น โดยที่ 1 buffer จะกินหน่วยความจำ 256 KB และถูกแบ่งออกเป็น 4 bank ใน 1 bank จะมีอยู่ 64 line แต่ละ line จะมี 512 จุด แต่ละ bank เราสามารถเลือกได้โดยใช้ command register (ดูวิธีการ set bank ได้ในหนังสือ *Eyegrabber*)

Convolution

ขั้นตอนนี้จะนำเอาข้อมูลที่เก็บไว้ในแผ่น disk นำมาทำการประมวลผล โดยใช้โปรแกรม `vision.exe` ซึ่งโปรแกรมนี้อาจทำหน้าที่ดึงข้อมูลภาพข้างซ้าย และภาพข้างขวามาทำการ convolution ด้วย LoG Mask (ดังแสดงในบทที่ 6 ในหัวข้อ 6.7 บทส่งท้าย(A) หน้า 6-13) กับข้อมูลภาพในรูปแบบ 8-2 และ 8-3 ซึ่งเมื่อผ่านการ convolution แล้วจะได้ผลดังรูปที่ 8-5 และ 8-6 ข้อมูลที่ได้จะมีเครื่องหมายเป็น "+", "-", และ "0" ซึ่งได้แสดงไว้เป็นสีคือ "+" = สีขาว "-" = สีชมพู และ "0" = สีดำ ศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 8-2 ภาพกล้องที่รับมาจากกล้อง CCD

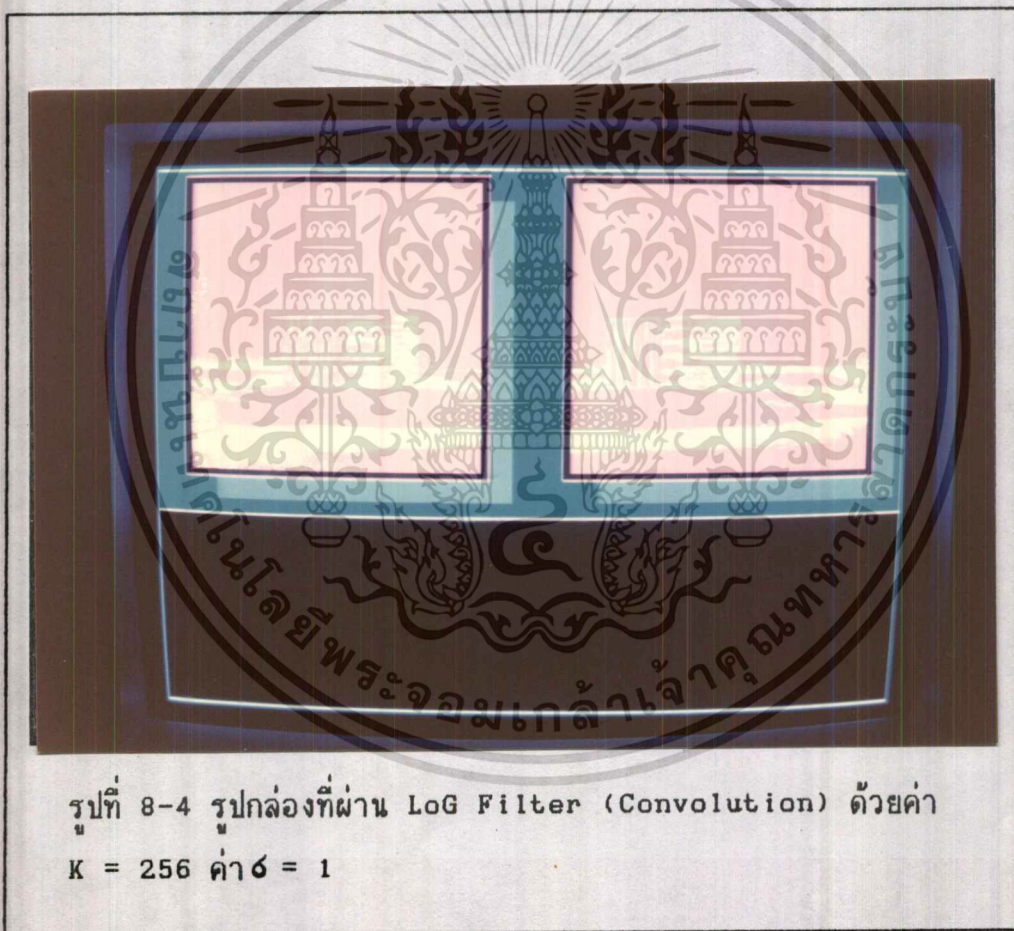


รูปที่ 8-3 ภาพรูปตัวอักษรที่รับมาจากกล้อง CCD

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ทางการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

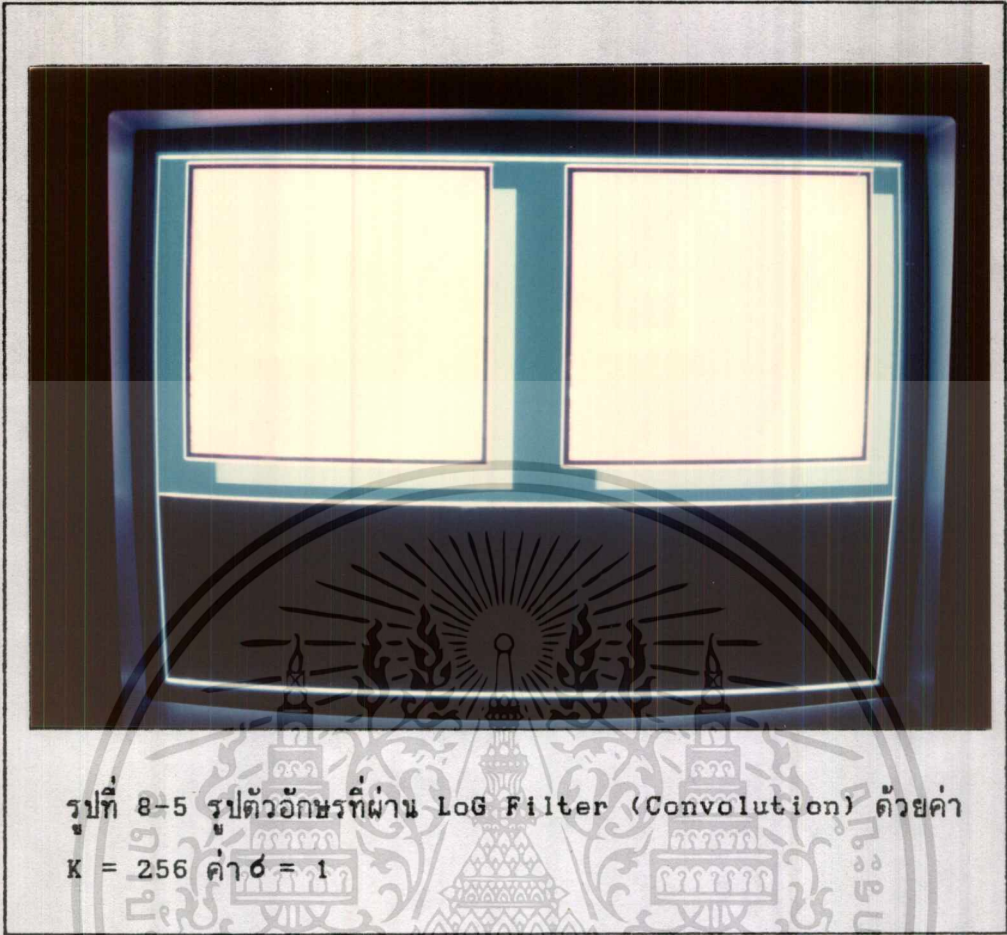
Zero-Crossing

ขั้นตอนนี้ยังคงเป็นขั้นตอนที่ต่อเนื่องมาจากการทำ Convolution และยังคงใช้โปรแกรม vision.exe ซึ่งในโปรแกรมนี้ ยังมี module การหาขอบภาพโดยวิธี zero-crossing ซึ่งใช้ Predicate of edge algorithm ในการคำนวณหาขอบภาพและทิศทางของขอบภาพ เมื่อผ่าน algorithm นี้แล้ว ข้อมูลภาพจะถูก encode ด้วยทิศทางของขอบภาพใน 4 บิตท้ายของข้อมูลภาพ รูปที่ 8-6 และ 8-7 (ทางซ้ายของภาพ) แสดงข้อมูลที่ผ่าน Predicate of edge algorithm แล้ว จะแสดงเป็นสีต่างๆ กัน 8 สี ซึ่งบ่งบอกถึงทิศทางของขอบภาพ



รูปที่ 8-4 รูปกล่องที่ผ่าน LoG Filter (Convolution) ด้วยค่า $K = 256$ ค่า $\sigma = 1$

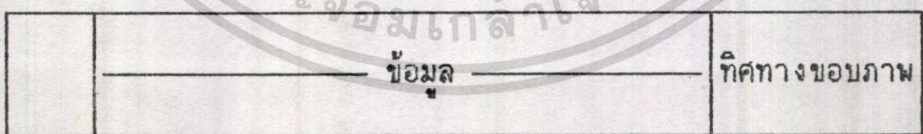
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 8-5 รูปตัวอักษรที่ผ่าน LoG Filter (Convolution) ด้วยค่า $K = 256$ ค่า $\sigma = 1$

ข้อมูลที่ผ่าน LoG Filter และ Predicate of edge algorithm จะมีรูปแบบการเก็บข้อมูลเป็นรูปข้างล่างนี้

15 14 4 3 0



└─ เครื่องหมาย (Signed Bit)

ซึ่งข้อมูลที่ได้นี้จะนำไปใช้ในการหาจุดที่สอดคล้องกันของภาพข้างซ้าย และภาพข้างขวา โดยใช้เงื่อนไขในการ matching ดังนี้

- 1) เครื่องหมาย (Signed Bit) ของภาพข้างซ้ายและภาพข้างขวาจะต้องเหมือนกัน
- 2) ทิศทางของขอบภาพของภาพข้างซ้ายและภาพข้างขวาจะต้องเหมือนกัน
- 3) ผลต่างของข้อมูล (data) ระหว่างภาพข้างซ้ายและภาพข้างขวาจะต้องมีค่าน้อย

เอกสารนี้เป็นเอกสารของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง หากมีข้อผิดพลาดประการใดขออภัยเป็นอย่างสูง และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

และในการ Matching เพื่อค้นหาตำแหน่งของข้อมูลที่สอดคล้องกัน (match) ได้โดยการสแกนหาข้อมูลภายใน "window" ขนาดเล็กเช่น ขนาด 21 x 7 , 41 x 7 หรือ 61 x 7 เป็นต้น ดังแสดงไว้ในรูป 8-6 และ 8-7 (ภาพทางขวา) ซึ่งข้อมูลที่ได้หลังจากผ่านการ Matching แล้ว จะถูกนำไปหาจุดที่สอดคล้องกัน เพื่อคำนวณหาระยะทางระหว่างเลนส์ของกล้องและวัตถุโดยใช้สมการในบทที่ 7



รูปที่ 8-6 รูปกล้องทางซ้ายแสดงทิศทางของขอบภาพด้วยสีต่างๆ 8 สี รูปกล้องทางขวามือแสดงจุดที่สามารถ match กันได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 8-7 รูปตัวอักษรทางซ้ายแสดงทิศทางของขอบภาพด้วยสีต่างๆ 8 สี
รูปตัวอักษรทางขวามือแสดงจุดที่สามารถ match กันได้

Make matching file

เป็นการ mark จุดพิกัด x, y ที่สามารถ match กันได้ แล้วเก็บเอาไว้ในไฟล์ .M เพื่อใช้ในการหาจุดที่สามารถ match กันได้ด้วยการเลื่อน mouse โปรแกรมที่ทำหน้าที่ดังกล่าวนี้ คือ markmeth.exe

Find distance from lens to object

ขั้นตอนนี้เป็นกรนำภาพที่หาขอบภาพเรียบร้อยแล้วมาคำนวณหาระยะทางจากเลนส์กล้องถึงตัววัตถุ โดยจะใช้ภาพข้างซ้ายเป็นต้นแบบ (master) จากนั้นจะใช้ mouse ทำการเลื่อนหาจุด x, y ที่ต้องการทราบระยะทาง เมื่ออัด x, y ข้างซ้าย จะมีพิกัด x, y ทางขวาที่จุดใด จากนั้นจึงนำไปเข้าสู่ตรดงแสดงไว้ข้างล่างนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นอนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$Z = R \sin \theta$$

$$d \sqrt{f^2 + x_1^2 + x_2^2}$$

$$R = \frac{d \sqrt{f^2 + x_1^2 + x_2^2}}{|x_1 + x_2|}$$

$$\theta = \tan^{-1} f/x_1$$

เมื่อเราเลือกโชค่า $d = 10,8,6$

$$f = 7 \text{ mm.}$$

$$x_1 = K \times \text{จำนวนพิกเซลที่นับได้จากจอภาพซ้าย}$$

$$x_2 = K \times \text{จำนวนพิกเซลที่นับได้จากจอภาพขวา}$$

$$K = 0.009 \text{ mm/pixel}$$

วิธีคำนวณหาค่า K

เราจะหาค่า K ได้จากการเทียบอัตราส่วน ระหว่างพิกเซลที่อิมเมจกับพิกเซลที่หน้าจอ โดยเทียบว่า 1 pixel ในจอภาพนั้น เป็นกี่ mm ในอิมเมจเพลน (image plane) จะได้ว่า

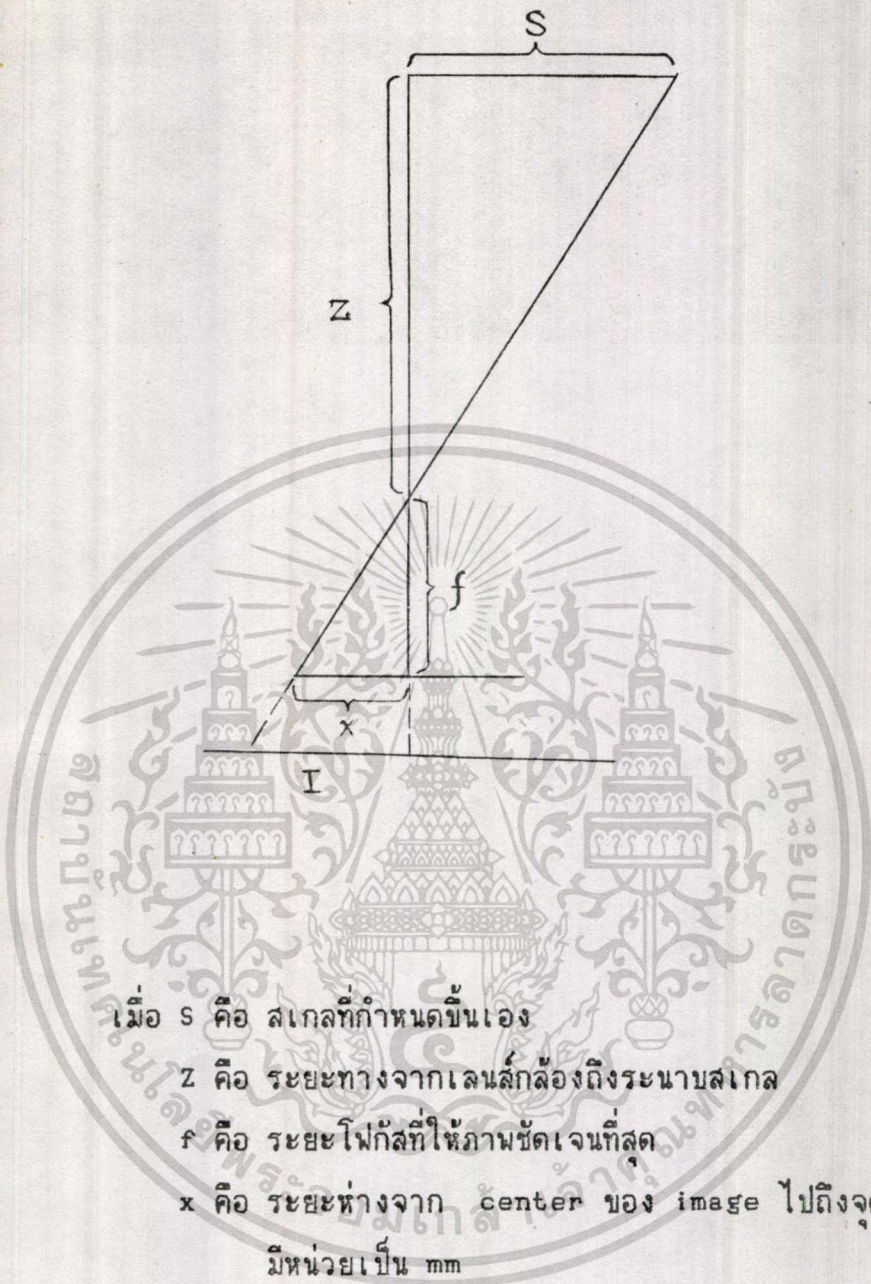
$$x = K \times I \quad ; \quad K < 0$$

เมื่อ K คือ อัตราส่วนที่ได้จากการเทียบสเกล

I คือ จำนวนพิกเซลที่จอภาพ

x คือ ค่าจาก Center ของอิมเมจไปยังจุด P ที่ตกกระทบ มีหน่วยเป็น mm

เราสามารถหาค่า x โดยการเทียบสเกล ดังรูปข้างล่าง



เมื่อ s คือ สเกลที่กำหนดขึ้นเอง
 z คือ ระยะทางจากเลนส์กล้องถึงระนาบสเกล
 f คือ ระยะโฟกัสที่ให้ภาพชัดดีที่สุด
 x คือ ระยะห่างจาก center ของ image ไปถึงจุด P ที่ตกกระทบ
 มีหน่วยเป็น mm

ซึ่งจากรูปข้างบนจะเป็นสามเหลี่ยมคล้าย ซึ่งเขียนเป็นสมการได้ดังนี้

$$\frac{x}{f} = \frac{s}{z} \dots\dots\dots (A)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\therefore x = \frac{S \times f}{z} \dots\dots\dots (B)$$

เมื่อเราได้ค่า x จากสมการ (B) เราก็สามารถคำนวณหาค่า K ได้ดังนี้

$$K = \frac{x}{I}$$

โดย I จะเป็นค่าที่ได้จากการนับ pixel บนจอภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยมนำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ผลการทดลอง

$$d = 10 \text{ cm.}$$

$$f = 7 \text{ mm.}$$

$$k = 0.009 \text{ mm./pixel}$$

L(x)	R(x)	Different (pixel)	Z(วัดได้) (cm.)	Z(คำนวณ) (cm.)	+ % ผิดพลาด
26	-59	85	100	91.76	8.24
21	-64	85	100.7	91.81	8.82
14	-70	84	101.4	92.97	8.31
9	-75	84	102.3	93.02	9.07
3	-80	83	103.2	94.20	8.72
-2	-84	82	104.05	95.40	8.31
-7	-89	82	104.9	95.47	8.98
-12	-94	82	105.65	95.54	9.56
-17	-97	80	106.4	97.98	7.91
-23	-103	80	107.1	98.07	8.43

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$d = 8 \text{ cm.}$$

$$f = 7 \text{ mm.}$$

$$k = 0.009 \text{ mm./pixel}$$

L(x)	R(x)	Different (pixel)	Z(วัดได้) (cm.)	Z(คำนวณ) (cm.)	+ % ผิดพลาด
14	-49	63	100	98.96	1.04
10	-53	63	100.7	99.00	1.68
3	-58	61	101.4	102.28	0.87
-2	-63	61	102.3	102.34	0.04
-8	-68	60	103.2	104.10	0.87
-13	-73	60	104.05	104.16	0.11
-18	-78	60	104.9	104.22	0.65
-23	-82	59	105.65	106.05	0.37
-29	-87	58	106.4	107.95	1.46
-33	-91	58	107.1	108.01	0.85

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$d = 6 \text{ cm.}$$

$$f = 7 \text{ mm.}$$

$$k = 0.009 \text{ mm./pixel}$$

L(x)	R(x)	Different (pixel)	Z(วัดได้) (cm.)	Z(คำนวณ) (cm.)	+ % ผิดพลาด
5	-41	46	100	101.59	1.59
1	-44	45	100.7	103.86	3.14
-6	-51	45	101.4	103.93	2.49
-10	-55	44	102.3	103.96	1.62
-17	-61	43	103.2	106.38	3.08
-22	-65	43	104.05	108.90	4.66
-27	-70	43	104.9	108.96	3.87
-32	-75	43	105.65	109.02	3.18
-37	-80	43	106.4	109.10	2.53
-41	-84	43	107.1	109.15	1.91

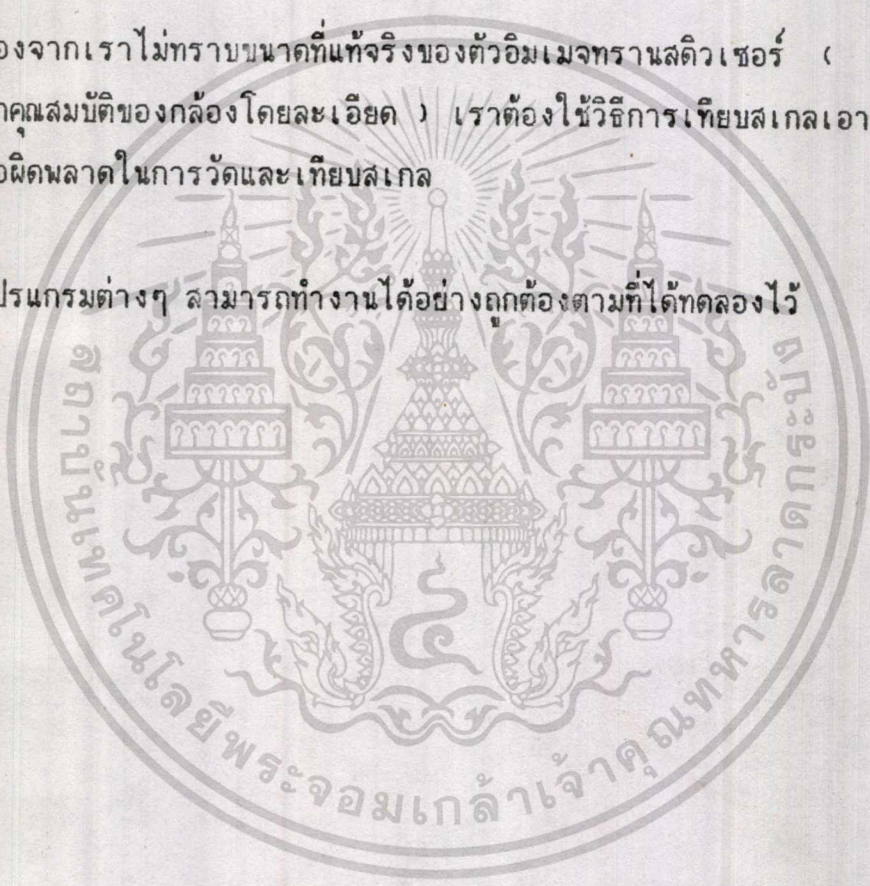
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สรุปผลการทดลอง

จากตารางผลการทดลองจะเห็นได้ว่ายังมีความผิดพลาดอยู่ทั้งนี้เราขอที่จะสรุปเป็นข้อๆได้ ดังนี้

- 1) อุปกรณ์ที่ใช้ในการทดลอง เช่น กล้องที่ใช้ในการทดลองออกแบบมาไว้ใช้งานที่ไม่ต้องการความแม่นยำ (เป็นกล้องที่ใช้เก็บภาพวิดีโอธรรมดา)
- 2) การทดลองเราใช้กล้องเพียงตัวเดียวในการเก็บภาพแต่ละครั้ง เพื่อนำมาทำการประมวลผล ซึ่งทำให้ต้องมีการปรับแต่งการวางตำแหน่งของกล้องให้ถูกต้องที่สุดเท่าที่จะทำได้
- 3) เนื่องจากเราไม่ทราบขนาดที่แท้จริงของตัวอิมเมจทรานสดิวเซอร์ (คู่มือกล้องไม่ได้บอกคุณสมบัติของกล้องโดยละเอียด) เราต้องใช้วิธีการเทียบสเกลเอาเอง ซึ่งอาจจะมีข้อผิดพลาดในการวัดและเทียบสเกล

ส่วนโปรแกรมต่างๆ สามารถทำงานได้อย่างถูกต้องตามที่ได้ทดลองไว้



กิติกรรมประกาศ

ขอขอบพระคุณ บิดา มารดา ที่ได้ให้การสนับสนุนการศึกษามาโดยตลอด และ ปรียญาตินิพนธ์ฉบับนี้ ได้สำเร็จลงด้วยดี ก็ด้วยความช่วยเหลือจากอาจารย์ และเพื่อนๆ ขอขอบพระคุณ รศ.ดร.นุศักดิ์ ชิวสุวิทย์ ที่ได้ให้คำปรึกษาและแนะนำตลอดจนได้อำนวยความสะดวกในการใช้เครื่องมือเครื่องมือนในการทดลอง ขอขอบพระคุณ อาจารย์ เกษตร์ ศิริสันติสัมฤทธิ์ ที่ได้ให้คำปรึกษาและแนะนำในการแก้ไขปัญหิต่างๆ จนการทำปรียญาตินิพนธ์ได้สำเร็จลง ขอขอบพระคุณ อาจารย์ทุกท่านในภาคเทคโนโลยี การวัดคุมอุตสาหกรรม ที่มีส่วนร่วมในข้อเสนอแนะต่างๆ ขอขอบคุณ พี่ๆ และ เพื่อนๆ ที่ได้ช่วยแก้ไขปัญหิต่างๆ และให้กำลังใจในการทำงาน



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หนังสืออ้างอิง

- [1] Adrian Low, "INTRODUCTORY COMPUTER VISION AND IMAGE PROCESSING", McGRAW-HILL Book Company, 1991.
- [2] Louis J. Galbati, Jr., "MACHINE VISION AND DIGITAL IMAGE PROCESSING FUNDAMENTAL", Prentice Hall International, Inc., 1990.
- [3] Rafael C. Gonzales, "DIGITAL IMAGE PROCESSING", Second Edition, Addison-Wesley Publishing Company, 1987.
- [4] Andrew C. Staugaard, Jr. " ROBOTICS AND AI : AN INTRODUCTION TO APPLIED MACHINE INTELLIGENCE ", Prentice-Hall.
- [5] Patrick Henry Winston , "ARTIFICIAL INTELLIGENCE", Second Edition
- [6] Anil K. Jain, "FUNDAMENTALS OF DIGITAL IMAGE PROCESSING", Prentice-Hall International Inc, 1989.
- [7] Andres Huertas and Gerard Medioni, " DETECTION OF INTENSITY - CHANGES WITH SUBPIXEL ACCURACY USING LAPLACIAN-GAUSSIAN MASKS ", IEEE Trans. Pattern Anal. Machine Intell., VOL.PAMI-8.NO.5 Sept. 1986.
- [8] W. Eric L. Grimson , " COMPUTATIONAL EXPERIMENTS WITH A FEATURE BASED STEREO ALGORITHM ", IEEE Trans. Pattern Anal. Machine Intell., VOL. PAMI-7 NO.1, Jan 1985.
- [9] Mun K. Leung and Thomas S. Huang, " MATCHING IMAGE POINT FEATURES OVER SPACE AND TIME ", Coordinated Science Lab. University of Illinois.
- [10] Stephen T. Barnard, " DISPARITY ANALYSIS OF IMAGE " , IEEE Trans. Pattern Anal. Machine Intell., VOL. PAMI-2 NO.4 July 1980.
- [11] "THE KING EYEGRABBER IMAGE ADAPTER" (MANUAL)
- [12] "BASIC IMAGING SOFTWARE PACKAGE FOR GENSTAR SYSTEMS, INC.", 1988.
- [13] "OPERATING INSTRUCTION FOR VIDEO CAMERA RECORDER", Sony, 1990.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้