ปีการศึกษา 2533

ระบบการจดจำลายนิ้วมือ

FINGERPRINT RECOGNITION SYSTEM

โดย

นาย กิตติพงษ์ ศิลปะ

นาย ชัยโรจน์ พัฒนพิทูรย์

นาย ชุมพล สุวิชาโน

นาย ธีรยุทธ์ อมรพุทธพล

อาจารย์ที่ปรึกษา

อาจารย์ เกษตร์ ศิริสันติสัมฤทธิ์

เลขหมู่ T 330861 ก3

เลขทะเบียน 02791

วัน, เดือน, ปี 7 สิงหาคม 34

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปริญญานิพนธ์ปีการศึกษา 2533

ภาควิชา เทคโนโลยีการวัดคุมทางอุตสาหกรรม

คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง ระบบการจดจำลายนิ้วมือ (FINGERPRINT RECOGNITION SYSTEM)

ผู้จัดทำ

1. นาย กิตติพงษ์ ศิลปะ 30-1018
2. นาย ชัยโรจน์ พัฒนพิฑูรย์ 30-1057
3. นาย ชุมพล สุวิชาโน 30-1061
4. นาย ธีรยุทธ์ อมรพุทธพล 30-1099

.....อาจารย์ที่ปรึกษา

(อาจารย์ เกษตร์ ศิริสันติสัมฤทธิ์)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ระบบการจดจำลายนิ้วมือ

กิตติพงษ์ ศิลปะ
ชัยโรจน์ พัฒนพิฑูรย์
ชุมพล สุวิชาโน
ธีรยุทธ์ อมรพุทธพล
เกษตร์ ศิริสันติสัมฤทธิ์ อาจารย์ที่ปรึกษา
ปีการศึกษา 2533

บทคัดย่อ

จุดประสงค์ของวิทยานิพนธ์เล่มนี้ คือ ต้องการแสดงถึงวิธีการที่จะนำเอาการประมวลผลข้อมูลภาพ มาใช้ประยุกต์ในงานด้าน การจดจำลายนิ้วมือ

เทคนิคที่ใช้ในระบบการจดจำลายนิ้วมือ ก็คือ **ระบบทรี** (Tree System) ซึ่งแบ่งได้เป็น 2 ประเภทใหญ่ๆ คือ **ภาษาทรี** (Tree Language) **และขบวนการจดจำแบบทรี** (Tree Automata) ภาษาทรี ใช้อธิบายลักษณะของลายนิ้วมือได้ ส่วนขบวนการจดจำแบบทรี ใช้เพื่อจดจำรูปแบบรหัสที่เราแปลงได้ ซึ่งจะมีประโยชน์มากในการใช้ค้นหา และเก็บรักษาข้อมูลที่มีจำนวนมากให้ลดลงได้ จึงทำให้เกิดความสะดวก ลดความสับสน และซับซ้อนลง โดยเฉพาะอย่างยิ่ง ยังสามารถลดเวลาในการค้นหาข้อมูลจำนวนมากได้อีกด้วย

ดังนั้น การวิจัย และพัฒนาระบบการจดจำลายนิ้วมือ จึงมีประโยชน์ต่องานปราบปรามอาชญากรรม และระบบการรักษาความปลอดภัยได้อีกด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

FINGERPRINT RECOGNITION SYSTEM

KITTIPONG SILPA

CHAIROJ PATTHANAPITOON

CHUMPOL SUWICHANO

TEERAYUT AMORNPUTTAPON

KASET SIRISANTISAMRID ADVISOR

ACADEMIC YEAR 1990.

ABSTRACT

The purpose of this thesis is to demostrate how to use the image processing for applying in fingerprint recognition system.

The technique used for fingerprint recognition system is tree system which can be composed of 2 categories - tree language and tree automata, the first is used to describe the fingerprint patterns and the set of tree automata is used to recognize the coded patterns. This system is very useful in finding and keeping the numerous informations to reduce the data, which is the conveniences, not more confused and complex, especially for reducing the time used.

So, in the research and developement the fingerprint recognition system is useful to criminal investigation and security systems.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

บทที่ 1 บทนำ.....1

1.1 พื้นฐานเบื้องต้น.....1

1.2 ระบบการแทนของข้อมูลภาพ.....2

1.3 องค์ประกอบของระบบประมวลภาพแบบดิจิตอล.....3

บทที่ 2 การทำให้ภาพชัดขึ้น.....7

2.1 วิธีในสเปทเชียลโดเมน.....7

2.2 วิธีในโดเมนเชิงความถี่.....8

บทที่ 3 การตรวจจับขอบของภาพ.....14

3.1 พื้นฐานเบื้องต้น.....14

3.2 การหาจุดเปลี่ยนแปลงที่ผ่านศูนย์ของฟิลเตอร์แบบ LoG...15

3.3 เทคนิคการแยกตัวปฏิบัติการแบบ LoG.....16

3.4 การทำนายทิศทางของภาพ.....19

3.5 การทดลองและผลการทดลอง.....42

3.6 สรุป.....45

บทที่ 4 ระบบการจดจำลายนิ้วมือ.....47

4.1 พื้นฐานเบื้องต้น.....47

4.2 ระบบจดจำลายนิ้วมือ.....49

4.3 การทดลองและผลการทดลอง.....63

4.4 สรุปผลการทดลองและแนวทางในการพัฒนา.....71

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5 ระบบเกี่ยวกับการมองเห็น.....73

5.1 รูปแบบของอิมเมจ.....73

5.2 การมองแบบสเตอริโอ.....75

5.3 คำแนะนำสำหรับการดูแบบสเตอริโอ.....81

5.4 การดูแบบสเตอริโอ.....82

ภาคผนวก

กิตติกรรมประกาศ.....

หนังสืออ้างอิง.....

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1

บทนำ

1.1 พื้นฐานเบื้องต้น [1]

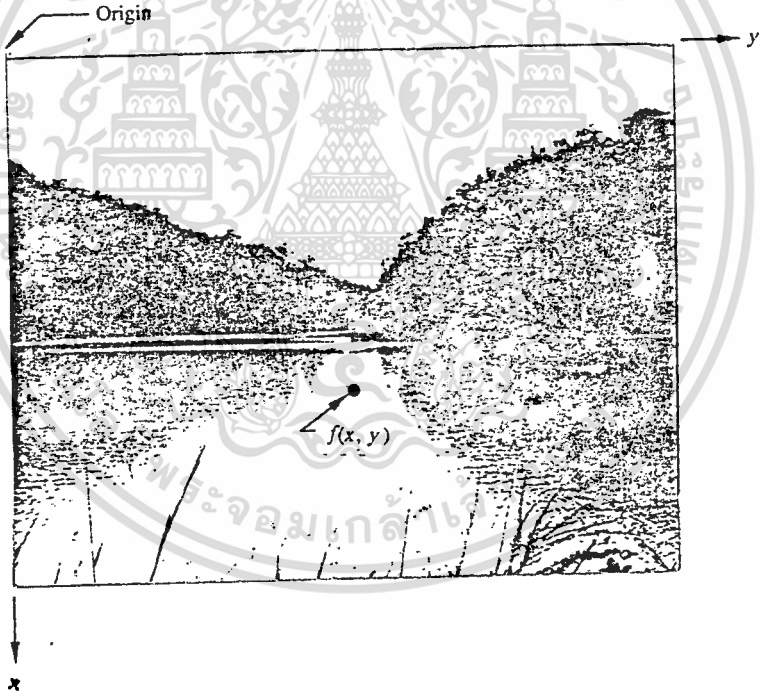
สิ่งที่น่าสนใจในวิธีการประมวลข้อมูลภาพแบบดิจิตอล มี 2 ประการ คือ การประมวลข้อมูลภาพเพื่อความเข้าใจของมนุษย์ และการประมวลข้อมูลภาพ เพื่อความเข้าใจของเครื่องจักรอัตโนมัติ การประยุกต์ในประการแรกนี้ ทำขึ้นเพื่อปรับปรุงภาพที่ตีพิมพ์ลงในหนังสือพิมพ์ที่ส่งโดยสายเคเบิ้ลใต้น้ำระหว่างลอนดอน และนิวยอร์ค ในสมัยต้นทศวรรษ 1920 ทำให้สามารถลดระยะเวลาในการส่งจากนานกว่า 1 สัปดาห์ ให้เหลือเพียงระยะเวลาน้อยกว่า 3 ชั่วโมง ส่วนการประยุกต์ในประการที่ 2 ได้เริ่มปรับปรุงหลังจากการส่งข้อมูลภาพแบบดิจิตอลดำเนินมาแล้ว 35 ปี โดยใช้เครื่องคอมพิวเตอร์ที่มีขนาดใหญ่ เพื่อที่จะนำไปสู่แนวความคิดเกี่ยวกับความเข้าใจในข้อมูลภาพการปรับปรุงข้อมูลภาพ เพื่อความเข้าใจของเครื่องจักรอัตโนมัต ได้เริ่มต้นขึ้นที่ห้องทดลอง Jet Propulsion (Pasadena ,California) ในปี 1964 เมื่อภาพของดวงจันทร์ที่ส่งจากยานอวกาศแรงเยอรส์ 7 (Rangers 7) ถูกแก้ไขโดยเครื่องคอมพิวเตอร์ เพื่อที่จะลดความเพี้ยนของภาพ เมื่ออยู่ในกล้องโทรทัศน์ ซึ่งนับเป็นที่มาของการนำไปประยุกต์ใช้ประโยชน์ในหลายๆด้าน เช่น งานในวงการแพทย์ งานในวงการวิศวกรรม งานในวงการอุตสาหกรรม งานในวงการทหาร งานในวงการตำรวจ หรือแม้กระทั่งงานที่เกี่ยวกับระบบการรักษาความปลอดภัย ดังที่จะนำเสนอในบทที่ 4 โดยในที่นี้จะเสนอพื้นฐานเบื้องต้นของการประมวลข้อมูลภาพแบบดิจิตอลโดยทั่วไปเป็นอันดับแรก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.2 ระบบการแทนของข้อมูลภาพ [1]

(Digital Image Representation)

ในวิทยานิพนธ์เล่มนี้ จะอ้างถึงระบบภาพโดยทั่วไป ให้แทนด้วยฟังก์ชันของความเข้มแสงในระนาบ 2 มิติ : f(x,y) โดยที่ x และ y แทนคู่ลำดับที่เกิดขึ้นที่ภาพจริง และค่าของฟังก์ชัน f ณ จุด (x,y) ใดๆ จะเป็นสัดส่วนโดยตรงกับความสว่าง หรือระดับสีเทา (grey level) ของภาพที่จุดนั้นๆ ตามตัวอย่างที่ได้แสดงในรูป 1.1



รูปที่ 1.1 แนวแกน x,y ที่ใช้ในแทนในระบบภาพดิจิตอล

สำหรับข้อมูลภาพแบบดิจิตอล f(x,y) จะมีลักษณะที่ไม่ต่อเนื่องทั้งในระนาบสเปทเชียล (spatial) ดังนั้นจึงพิจารณาให้ข้อมูลภาพแบบดิจิตอลแทนด้วยเมตริกซ์ๆหนึ่ง ซึ่งมีแถว และหลัก ที่มีลักษณะเป็นเอกลักษณ์ และมีความสัมพันธ์กับค่าสมาชิกในแถว และหลักที่แทนด้วยระดับสีเทาของภาพที่จุดนั้นๆ

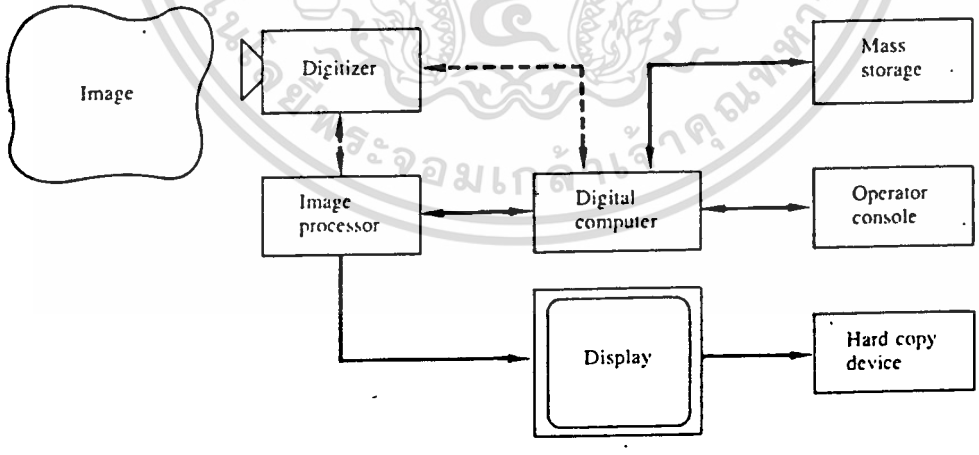
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ถึงแม้ว่าขนาดของข้อมูลภาพแบบดิจิตอล จะแปรเปลี่ยนไปตามการประยุกต์การใช้งาน แต่จะมีข้อได้เปรียบที่ดี ถ้ากำหนดให้มีลักษณะเฉพาะเป็นอะเรย์แบบสมมาตร (square array) และจำนวนของระดับสีเทาเป็นเลขจำนวนเต็มกำลังสอง เช่น ภาพที่ได้จากโทรทัศน์โมโนโครม มีขนาด 512x512 พิเซล และมีจำนวน 128 ระดับสีเทา เป็นต้น

1.3 องค์ประกอบของระบบประมวลข้อมูลภาพแบบดิจิตอล

(Elements of A Digital Image Processing System)

องค์ประกอบพื้นฐานที่จำเป็นโดยทั่วๆไป ในระบบการประมวลข้อมูลภาพแบบดิจิตอลใดๆ ได้แสดงในรูป 1.2 การทำงานของแต่ละบล็อคในรูป 1.2 นี้ สามารถอธิบายได้โดยย่อดังนี้



รูปที่ 1.2 แสดงองค์ประกอบของระบบการประมวลผลแบบดิจิตอล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.3.1 เครื่องประมวลข้อมูลภาพ (Image Processor)

เครื่องประมวลข้อมูลภาพแบบดิจิตอล เป็นหัวใจของระบบการประมวลข้อมูลภาพ เครื่องประมวลข้อมูลภาพ ประกอบด้วยชุดของฮาร์ดแวร์ที่มี 4 ฟังก์ชั่นพื้นฐาน คือ การหยุดภาพนิ่ง (image acquisition) การเก็บข้อมูลภาพ (storage) การประมวลภาพในระดับต่ำ(เร็ว) (low-level (fast) processing) และการแสดงผล (display) โดยทั่วไปโมดูลของการหยุดภาพนิ่ง จะมีสัญญาณโทรทัศน์เป็นหน่วยอินพุท และทำการเปลี่ยนสัญญาณนี้ให้เป็นข้อมูลแบบดิจิตอล เครื่องประมวลข้อมูลภาพสมัยใหม่ส่วนมากใช้ความสามารถของการดิจิไตซ์ภาพโทรทัศน์ในเวลา 1 เฟรม (one frame-time) เช่น ในเวลา 1/30 วินาที จึงมักเรียกว่า เครื่องจับภาพนิ่ง (frame grabber)

โมดูลของการเก็บภาพ ที่เรียกว่า เฟรมบัฟเฟอร์ (frame buffer) จะทำหน้าที่เก็บรักษาหน่วยความจำของภาพแบบดิจิตอล โดยทั่วไปจะมีเครื่องมือ 2-3 ชุดที่รวมกันเข้าไว้ในเครื่องประมวลข้อมูลภาพ คุณสมบัติการหยุดภาพของโมดูลการเก็บภาพประกอบด้วยหน่วยความจำที่สามารถอ่านข้อมูลจากอัตราของโทรทัศน์ มักเป็น 30 ภาพต่อวินาที ลักษณะนี้เป็นไปตามการทำงานของโมดูลการหยุดภาพนิ่งที่ประกอบไปด้วยภาพที่สมบูรณ์ในการเก็บภาพที่เร็วเท่าที่ภาพถูกจับให้นิ่ง หน่วยความจำสามารถเปลี่ยนให้เป็นแอดเดรสของโทรทัศน์ได้โดยชุดแสดงผลของภาพที่จอมอนิเตอร์

เครื่องประมวลข้อมูลภาพ ยังมีหน้าที่ในระดับต่ำอีก (low-level function) เช่น การประมวลผลทางคณิตศาสตร์ และทางตรรกศาสตร์ หรือที่เรียกว่า หน่วยตรรกเลขคณิต (Arithmetic-Logic Unit : ALU) โดยถูกออกแบบพิเศษ เพื่อที่จะเพิ่มความเร็วในการประมวลผลแบบขนาน ส่วนโมดูลแสดงผล (display module) ทำหน้าที่อ่านหน่วยความจำของภาพ เปลี่ยนข้อมูลของภาพที่เป็นแบบดิจิตอลให้เป็นสัญญาณวิดิโอแบบอนาลอก และแสดงผลสัญญาณที่แปลงได้นั้น ให้ออกทาง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

มอนิเตอร์โทรทัศน์ หรืออุปกรณ์วีดิโอชนิดอื่น

1.3.3 เครื่องคอมพิวเตอร์แบบดิจิตอล (Digital Computer)

ถึงแม้ว่า เครื่องประมวลข้อมูลภาพ อาจเทียบได้กับความสามารถในการประมวลผลภายใน ดังที่ได้กล่าวมาแล้วในหัวข้อ 1.3.1 แต่ว่าระดับของการประมวลผลยังอยู่ในระดับค่อนข้างต่ำในด้านความทันสมัย ดังนั้นโดยทั่วไปพบว่า เครื่องประมวลข้อมูลภาพนี้จะถูกต่อเชื่อม (interface) กับเครื่องคอมพิวเตอร์ ระบบคอมพิวเตอร์ที่ใช้สำหรับการประมวลข้อมูลภาพมีตั้งแต่ อุปกรณ์จำพวกไมโครโปรเซสเซอร์ จนถึงระบบคอมพิวเตอร์ขนาดใหญ่ ที่มีความสามารถในการประมวลข้อมูลภาพ ที่มีอะเรย์ใหญ่ๆ ด้วย พารามิเตอร์สำคัญ ที่มีอิทธิพลต่อโครงสร้างของเครื่องคอมพิวเตอร์ ที่ใช้สำหรับการประมวลข้อมูลภาพ คือ การเพิ่มการประยุกต์ใช้งาน และความต้องการขนาดของข้อมูลในการประมวลผล สำหรับการประยุกต์ใช้ที่พบกันทั่วไป จะเป็นเครื่องคอมพิวเตอร์ส่วนบุคคล หรือมินิคอมพิวเตอร์ แต่ถ้าต้องการใช้เกี่ยวข้องกับโปรแกรมที่มีขนาดใหญ่ในการพัฒนา หรือมีข้อมูลมาก ก็อาจใช้เครื่องคอมพิวเตอร์ระดับเมนเฟรมได้ ในกรณีนี้เครื่องคอมพิวเตอร์ที่มีหน่วยความจำของแอดเดรสมาก และมีความเร็วสูงก็ยิ่งได้เปรียบ

1.3.4 อุปกรณ์ที่ใช้ในการเก็บรักษาข้อมูล (Storage Devices)

ข้อมูลภาพแบบดิจิตอลที่ใช้มีขนาด 512x512 พิเซล แต่ละควอนไตซ์มีข้อมูลขนาด 8 บิท โดยมีสื่อที่ใช้ในการเก็บรักษาข้อมูลที่สำคัญ คือ แผ่นจานแม่เหล็ก (magnetic disks) เทปแม่เหล็ก (magnetic types) และแผ่นจานนำแสง (optical disks) แต่สื่อที่นิยมใช้ในการเก็บข้อมูลในระดับทั่วไป และใช้ในการวิจัยนี้ก็คือ แผ่นจานแม่เหล็ก ที่คุ้นเคยกันนั่นเอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.3.5 อุปกรณ์ที่ใช้ในการแสดงผล และบันทึกผล
(Display and Recording Devices)

จอโมโนโครม หรือจอโทรทัศน์สี เป็นอุปกรณ์ที่ใช้ในการแสดงผลที่สำคัญ และใช้กันมากในระบบการประมวลข้อมูลภาพสมัยใหม่ มอนิเตอร์จะถูกขับโดยเอาท์พุท ของชุดแสดงผลภาพในเครื่องประมวลข้อมูลภาพ สัญญาณที่ได้รับนี้สามารถเข้าสู่ระบบการบันทึกภาพที่อาจใช้ผลิตฮาร์ดก๊อปปี้ได้ (hard copy) เช่น สไลด์สี, รูปภาพ และแผ่นใส ในขณะที่กำลังดูภาพจากมอนิเตอร์ก็ได้ ส่วนสื่อที่ใช้ในการแสดงผล และบันทึกผลที่พบกันในทั่วไปก็คือ จอแบบหลอดภาพรังสีคาโธด (Cathod Ray Tube : CRT) และอุปกรณ์จำพวกเครื่องพิมพ์นั่นเอง

ในระบบของจอแบบหลอดภาพรังสีคาโธด ตำแหน่งของภาพทั้งในแนวฮอริเซนทอล (horizental) และในแนวเวอร์ติเคิล (verticle) ของแต่ละสมาชิกในอะเรย์ของภาพถูกเปลี่ยนให้เป็นโวลเตจ ซึ่งถูกยิงจากลำแสงอิเลคตรอนของหลอดภาพรังสีคาโธด โดยการขับลำแสงใน 2 มิติ ที่จำเป็นต่อการผลิตภาพเอาท์พุท ที่แต่ละจุดจะมีการยิงลำแสงอิเลคตรอน ความเข้มของลำแสงถูกเปลี่ยนโดยการใช้โวลเตจที่เป็นสัดส่วนโดยตรงต่อค่าของจุดที่มีความสัมพันธ์กันในอะเรย์เชิงเลข (numerical array) โดยการเปลี่ยนความเข้มจาก "0" ให้เป็นสีดำ และแทนความเข้มที่มากที่สุดด้วยสีขาว ผลการเปลี่ยนแปลงความเข้มแสงถูกบันทึกโดยกล้องถ่ายภาพแบบกราฟิก (photographic camera) ที่โฟกัสของจอภาพแบบหลอดภาพรังสีคาโธด

ส่วนอุปกรณ์แสดงผลภาพพิมพ์ที่ใช้กันในขั้นต้น จะมีความละเอียดต่ำ ซึ่งสามารถแสดงระดับสีเทาได้ด้วยการพิมพ์จำนวน และความเข้มของอักขระที่พิมพ์ในตำแหน่งนั้น โดยการเลือกอักขระที่เหมาะสมกับการกระจายในแบบระดับสีเทาที่ดี ก็จะทำให้เครื่องพิมพ์สามารถแสดงภาพระดับสีเทาได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

การทำให้ภาพชัดขึ้น

Image Enhancement

จุดประสงค์ของเทคนิคการทำให้ภาพชัด คือ ต้องทำภาพให้มีผลลัพธ์ที่เหมาะสมเฉพาะงานที่จะนำไปประยุกต์มากกว่าภาพต้นฉบับเดิม ซึ่งมีความสำคัญมาก เพราะว่าการใช้เทคนิคใดเทคนิคหนึ่งนั้น จะมีความเหมาะสมเฉพาะกับงานๆหนึ่งเท่านั้น ไม่สามารถนำเทคนิคเดียวกันนี้มาใช้ร่วมกันได้หมด การพิจารณาเทคนิคเหล่านี้ จะแบ่งออกเป็น 2 ประเภทใหญ่ๆ คือ วิธีในโดเมนเชิงความถี่ (frequency-domain method) และวิธีในโดเมนสเปทเชียล (spatial-domain method) เทคนิคแรกเป็นผลของการปรับปรุงการแปลงฟูเรียร์ (Fourier Transform) ของภาพ ส่วนอีกเทคนิคหนึ่งใช้การอ้างถึงระนาบภาพของตัวมันเอง และการปฏิบัติการโดยตรงกับพิเซลของภาพ

2.1 วิธีในโดเมนสเปทเชียล [1]

คำว่า สเปทเชียล มักหมายถึง การรวมกันของพิเซลของภาพ และวิธีในโดเมนสเปทเชียล คือ ขบวนการที่ปฏิบัติการโดยตรงกับพิเซลของภาพ ฟังก์ชันของการประมวลข้อมูลภาพในโดเมนสเปทเชียล อาจเขียนได้ว่า

$$g(x,y) = T[f(x,y)] \qquad \text{สมการที่ 2.1-1}$$

เมื่อ $f(x,y)$ คือ ภาพอินพุท (input image)

$g(x,y)$ คือ ผลของภาพที่ผ่านการปฏิบัติการแล้ว และ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

T คือ ตัวปฏิบัติการที่กระทำต่อ f(x,y) ซึ่งจะกำหนดขอบเขตของการปฏิบัติการเป็นย่านของ (x,y)

ดังที่แสดงในรูป 2.1 ถ้ากำหนดให้ w1,w2,...,w9 แทนสัมประสิทธิ์มาส์ก (mask) และพิจารณาที่รอบๆย่านนั้นเป็น 8 จุดรอบๆ (x,y)

w1 x-1,y-1	w2 x,y-1	w3 x+1,y-1
w4 x-1,y	w5 x,y	w6 x+1,y
w7 x-1,y+1	w8 x,y+1	w9 x+1,y+1

ซึ่งเราอาจกำหนดให้มีการปฏิบัติการดังนี้

$$T[f(x,y)] = w1*f(x-1,y-1) + w2*f(x,y-1) + w3*f(x+1,y-1) + w4*f(x-1,y) + w5*f(x,y) + w6*f(x+1,y) + w7*f(x-1,y+1) + w8*f(x,y+1) + w9*f(x+1,y+1)$$

2.2 วิธีในโดเมนเชิงความถี่ [1]

พื้นฐานของวิธีในโดเมนเชิงความถี่ คือ ทฤษฎีการทำคอนโวลูชัน ให้ g(x,y)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เป็นได้จากการทำคอนโวลูชันของภาพ f(x,y) และตัวปฏิบัติการที่ไม่เปลี่ยนตำแหน่ง (position-invarant operator) นั่นคือ

$$g(x,y) = h(x,y) * f(x,y) \qquad \text{สมการที่ 2.2-1}$$

จากทฤษฎีการทำคอนโวลูชัน จะได้ว่า

$$G(u,v) = H(u,v)F(u,v) \qquad \text{สมการที่ 2.2-2}$$

เมื่อ G,H และ F คือ การแปลงฟูเรียร์ของ g,h และ f ตามลำดับ ฟังก์ชัน H บางครั้งอาจเรียกว่า ทรานส์เฟอร์ฟังก์ชัน (transfer function) ดังนั้น จะได้ว่า

$$g(x,y) = L^{-1}[H(u,v)F(u,v)] \qquad \text{สมการที่ 2.2-3}$$

ซึ่งจะพบว่า g(x,y) คือ อินเวอร์สลาปลาส (inverse Laplace) ของ H(u,v) คูณกับ F(u,v) นั่นเอง

2.3 ฟิลเตอร์ [1]

ฟิลเตอร์ เป็นตัวกรองสัญญาณภาพขั้นต้น ให้ได้ภาพที่มีความคมชัดยิ่งขึ้น เพื่อที่จะนำภาพนั้นไปประยุกต์ใช้ได้ดียิ่งขึ้น ซึ่งมีมากมายหลายชนิด แต่ในที่นี้จะกล่าวเฉพาะที่นิยมใช้กันมากเท่านั้น

2.3.1 การเฉลี่ยค่ารอบย่าน (Neighborhood Average)

การเฉลี่ยค่ารอบย่าน เป็นเทคนิคที่ใช้กันในสเปทเชียลโดเมนโดยตรง สำหรับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสาร 027919 ไปใช้

การทำภาพให้ชัดยิ่งขึ้น ให้ภาพ f(x,y) ขนาด NxN และ g(x,y) เป็นผลที่ได้จากการปฏิบัติการของระดับสีเทาทุกๆ จุด (x,y) ได้รับจากการเฉลี่ยค่าระดับสีเทาของพิเซลในภาพ f(x,y) ซึ่งสามารถอธิบายได้ดังนี้

$$g(x,y) = 1/M \sum_{(n,m)} f(n,m) \qquad \text{สมการที่ } 2.3\text{-}1$$

สำหรับ x,y = 0,1,...,N-1 และ M คือ ผลรวมของจำนวนของจุดในย่านนั้นๆ

2.3.2 ฟิลเตอร์แบบมัธยฐาน (Median Filter)

สิ่งสำคัญประการหนึ่งของวิธีการที่ได้ในหัวข้อ 2.3.1 คือ การเกิดความเบลอร์ที่ขอบของภาพ และรายละเอียดสัญญาณรบกวนชัดเจน อาจแก้ไขได้โดยการใช้เทรสโชลด์ (threshold) ซึ่งโดยทั่วไปแล้วค่าเทรสโชลด์จะได้จากการลองผิดลองถูก (trial and error) แต่ถ้าใช้ฟิลเตอร์แบบมัธยฐานแทนระดับสีเทาของแต่ละพิเซล โดยการใช้ค่ามัธยฐานของระดับสีเทาในรอบย่านนั้นๆของพิเซลแทนการเฉลี่ย จะทำให้ภาพที่ได้มีประสิทธิภาพดีขึ้น

การหาค่ามัธยฐาน (M) คือ เซ็ตของค่าที่มีค่าในเซ็ตครึ่งหนึ่งมากกว่า M และอีกครึ่งหนึ่งน้อยกว่า M สามารถหาได้โดยทำการเรียงข้อมูลจากน้อยไปหามาก แล้วนำค่าที่อยู่ตรงกลางนั้น มาใช้แทนค่า M สมมติว่ามีย่านขนาด 3x3 มีข้อมูล (10, 20,20,20,15,20,20,25,100) จะสามารถเรียงข้อมูลใหม่ได้เป็น (10,15,20, 20,20,20,20,25,100) ดังนั้นค่ามัธยฐานคือ 20 นั่นเอง ความคิดที่มีอิทธิพลต่อฟังก์ชั่นของฟิลเตอร์แบบมัธยฐาน คือ จะทำให้จุดมีความเข้มที่แตกต่างมากกว่าย่านรอบๆของมัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3.3 ฟิลเตอร์แบบความถี่ต่ำผ่าน (low pass filter)

ขอบของภาพและสภาพที่คม เช่น สัญญาณรบกวนในระดับสีเทาของภาพมักจะมีส่วนของสัญญาณความถี่สูง ของการแปลงฟูเรียร์อย่างมาก ซึ่งจะทำให้เกิดการเบลอร์ ดังนั้น จึงหาวิธีการกำจัดเฉพาะช่วงความถี่สูงที่ไม่ต้องการออกไป จากหัวข้อ 2.2 เราได้

$$G(u,v) = H(u,v)F(u,v)$$

เมื่อ $F(u,v)$ คือ การแปลงฟูเรียร์ของภาพที่ต้องการให้เรียบ ปัญหาก็คือ ต้องเลือกฟังก์ชั่น $H(u,v)$ ที่ให้ได้ผลลัพธ์ $G(u,v)$ โดยสามารถกำจัดช่วงที่ความถี่สูงของ $F(u,v)$ ออกไป การแปลงอินเวอร์สลาปาสของ $G(u,v)$ จะทำให้ได้ว่า $g(x,y)$ เรียบขึ้น เพราะว่าส่วนประกอบความถี่สูงไม่ถูกกำจัดออกไป แต่ช่วงความถี่ต่ำจะสามารถผ่านไปได้ โดยปราศจากการลดทอนของสัญญาณ ซึ่งวิธีการนี้เรียกว่า ฟิลเตอร์แบบความถี่ต่ำผ่าน ฟังก์ชั่น $H(u,v)$ ถูกเรียกเป็น ทรานส์เฟอร์ฟังก์ชั่นของฟิลเตอร์ (Filter Transfer Function) ซึ่งในการคำนวนจะพิจารณาเฉพาะ $H(u,v)$ เท่านั้นฟิลเตอร์แบบความถี่ต่ำผ่านนี้ มีอยู่หลายชนิดด้วยกัน โดยที่แต่ฟิลเตอร์จะทำหน้าที่กรองความถี่สูงออกไป ให้เหลือเฉพาะความถี่ต่ำเท่านั้น เช่น

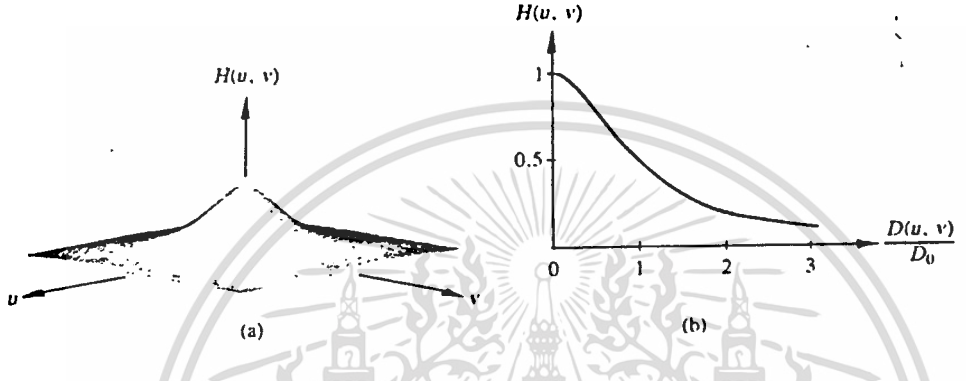
-ฟิลเตอร์แบบอุดมคติ (Ideal Filter) จะมี

$$H(u,v) = \begin{cases} 1 & \text{ถ้า } D(u,v) <= D_0 \\ 0 & \text{ถ้า } D(u,v) > D_0 \end{cases} \qquad \text{สมการที่ 2.3-2}$$

เมื่อ D_0 คือ ค่าจำเพาะที่ไม่เป็น 0 และ

$D(u,v)$ คือ ระยะทางจาก (u,v) ถึงจุดกำเนิดในระนาบความถี่ นั่นคือ $D(u,v) = (u^2 + v^2)^{0.5}$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



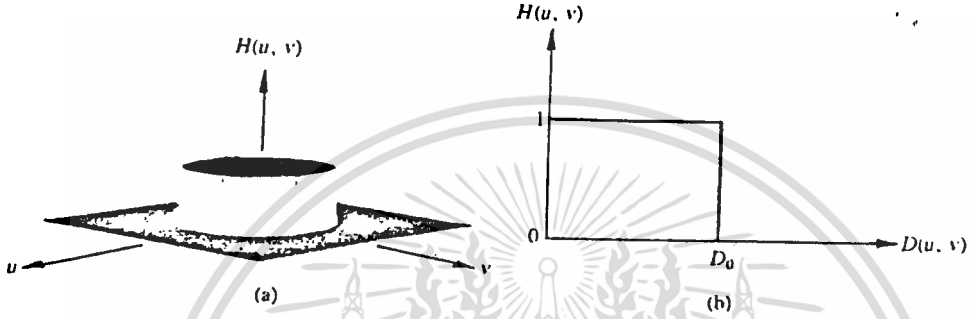
รูปที่ 2.2 (a) แสดงภาพฟิลเตอร์แบบความถี่ต่ำผ่าน (b) การตัดของฟิลเตอร์แบบบัตเตอร์เวิร์ท

- ฟิลเตอร์แบบบัตเตอร์เวิทธ์ (Butterworth Filter)

ฟังก์การแปลงของบัตเตอร์เวิทธ์ฟิลเตอร์ ของลำดับที่ n และมีความถี่ตัดออก (Cut-off Frequency locus) ที่ระยะทาง D_o จากจุดกำเนิดกำหนดโดยความสัมพันธ์

$$H(u,v) = 1/\{ 1+[D(u,v)/D_o]^{2n} \}$$ สมการที่ 2.3-3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.1 (a) แสดงภาพฟิลเตอร์แบบความถี่ต่ำผ่าน (b) การตัดของฟิลเตอร์ทางอุดมคติ

ความไม่เหมือนกันของฟิลเตอร์แบบความถี่ต่ำผ่านในทางอุดมคติ และฟิลเตอร์แบบความถี่ต่ำผ่านแบบบัตเตอร์เวิทธ์ คือ ไม่มีความต่อเนื่องที่ชัดเจนว่าตำแหน่งที่กำจัดความถี่ตัดออก ระหว่างความถี่ที่ยอมให้ผ่านกับความถี่ที่กรอง โดยทั่วไปมักกำหนดความถี่ตัดออกที่จุด ซึ่ง H(u,v) ได้ลดลงจากค่าที่มากที่สุด ในสมการ 2.3-3 จะพบว่า H(u,v) = 0.5 (ลดลง 50% จากค่าที่มากที่สุด) เมื่อ D(u,v) = D_0 แต่ค่าอื่นๆ โดยทั่วไปนิยมใช้ที่ 1/$\sqrt{2}$ ของค่าที่มากที่สุดของ H(u,v) ดังนั้นจากสมการ 2.3-3 จะได้ค่าที่ต้องการเมื่อ D(u,v) = D_0 คือ

$$H(u,v) = 1/\{1+[\sqrt{2}-1][D(u,v)/D_0]^{2n}\}$$

$$= 1/\{1+0.414[D(u,v)/D_0]^{2n}\}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

การตรวจจับขอบของภาพ

ในบทนี้ เป็นส่วนที่ได้ทำการวิจัย เกี่ยวกับทฤษฎีการตรวจจับขอบของภาพอีกวิธีหนึ่ง โดยใช้การเปลี่ยนแปลงที่ผ่านค่าศูนย์ ในหัวข้อ 3.1 และ 3.2 จะกล่าวถึงพื้นฐานเบื้องต้นเกี่ยวกับทฤษฎีการตรวจจับขอบของภาพ หัวข้อ 3.3 และ 3.4 เป็นเทคนิคที่ใช้ในการทดลองนี้ หัวข้อ 3.5 เป็นการทดลองและผลการทดลอง ส่วนในหัวข้อ 3.6 เป็นสรุปผลการทดลองที่ได้

3.1 พื้นฐานเบื้องต้น [2]

ขอบของภาพ คือ การเปลี่ยนแปลงระดับของความเข้มในภาพ

ขอบของภาพที่พบในภาพทั่วๆไป มีอยู่ 2 ลักษณะ

1. ขอบของภาพที่เกิดจากการเปลี่ยนแปลงความเข้มในธรรมชาติ ซึ่งเป็นขอบของภาพโดยแท้จริง ขอบของภาพในลักษณะนี้ ฟิลเตอร์ (filter) ที่เหมาะสมที่สุด คือ ตัวปฏิบัติการแบบอนุพันธ์ลำดับที่สอง (Second Derivative Operator) เพราะตัวปฏิบัติการแบบนี้จะไม่ขึ้นกับการหักเหของแสง การหาขอบของภาพโดยวิธีนี้ คือ การหาค่าที่เป็น 0 หรือ จุดเปลี่ยนแปลงที่ผ่านค่าศูนย์ (zero-crossing) จากการทำคอนโวลูชัน (convolution) ของ $\nabla^2 G(x,y) * I(x,y)$

2. การเปลี่ยนความเข้มในภาพที่เกิดความไม่ต่อเนื่อง เกิดจากการสะท้อนของแสง หรือเกิดจากขอบของภาพที่เกิดการสว่างขึ้น ซึ่งขอบของภาพลักษณะนี้จะใช้กับตัวปฏิบัติการแบบอนุพันธ์ลำดับที่สองได้ไม่ดีนัก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 การหาจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์ของฟิลเตอร์แบบ LoG [2]

โดยทั่วไปเราใช้ฟิลเตอร์แบบ LoG ของ Marr-Hildreth

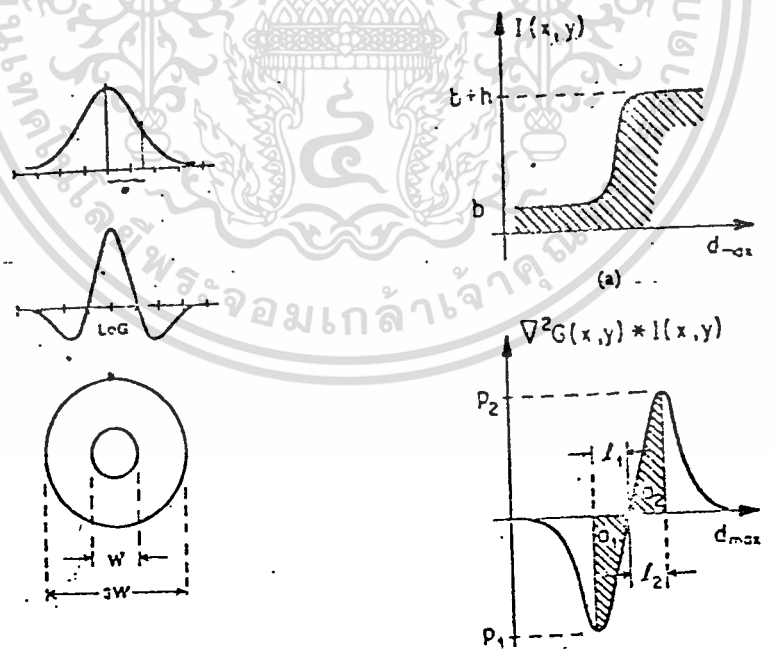
$$\nabla^2 G(x,y) = [1/2*pi*s^4]*[2-(x^2+y^2)/s^2]*exp[-(x^2+y^2)/2s^2]$$

สมการ 3.1

โดย s -ค่าเบี่ยงเบน (sigma) มีผลต่อความไว (sensitivity) ของภาพ

w -ความกว้างที่บริเวณส่วนกลางของตัวปฏิบัติการ (width of the central excitory region of the operator)

ซึ่ง $w = 2\sqrt{2}*s$



รูปที่ 3.1 แสดงลักษณะของ LoG มาส์ก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3 เทคนิคในการแยกตัวปฏิบัติการแบบ LoG [2]

(Separable Decomposition Of The LoG Operator)

เทคนิคที่กล่าวนี้ สามารถลดเวลาการคำนวนได้ ซึ่งทำได้โดยการแยกฟิลเตอร์แบบ LoG ให้เป็น 2 ส่วน ทำให้ใช้เวลาลดลง จาก M^2 เป็น 2M ได้ โดยวิธีการแยกทำได้ดังนี้

จาก

$$\nabla^2 G(x,y) = K[2-(x^2+y^2)/s^2]*\exp[-(x^2+y^2)/2s^2]$$

สมการ 3.2

โดยที่ K -ค่าคงที่สเกล (scale constant)

s -ค่าเบี่ยงเบน

จะได้ว่า

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

โดยที่ $h_{12}(x,y) = h_1(x)*h_2(y)$

$h_{21}(x,y) = h_2(x)*h_1(y)$

$h_1(a) = \sqrt{K}*[1-a^2/s^2]*\exp[-a^2/2s^2]$

$h_2(a) = \sqrt{K}*\exp[-a^2/2*s^2]$

โดยที่ a เป็นตัวแปรใดๆ

ค่า K ที่เลือกจะเป็นสัมประสิทธิ์สูงสุดของฟิลเตอร์ ผลรวมของสัมประสิทธิ์ของฟิลเตอร์มีค่าเป็น 0 ดังที่แสดงในตัวอย่าง การคำนวนจะลดเวลาลง โดยการแยกการทำคอนโวลูชันทั้ง 2 ด้าน เพื่อที่จะลดการเข้าถึงข้อมูล และยังได้เปรียบที่ฟิลเตอร์ที่ได้จะมีลักษณะสมมาตรทั้งแถว และหลัก ซึ่งจะลดจำนวนครั้งของการคูณโดยตัวประกอบอื่นๆ ให้เหลือ 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3.1 ตัวอย่างของการแยกตัวปฏิบัติการแบบ LoG

ตามตัวอย่างนี้จะอธิบายการแย๊กส่วนของฟิลเตอร์แบบ LoG ขนาดเล็กๆ 11x11 ที่มีค่าเบี่ยงเบนเท่ากับ $\sqrt{2}$ และค่าคงที่สเกลเท่ากับ 2116 เป็นฟิลเตอร์แถว (row filter) และฟิลเตอร์หลัก (column filter) ขนาด 1 มิติแล้วนำมาสร้างใหม่เป็นฟิลเตอร์แบบ LoG ที่มีขนาด 2 มิติ และเทียบกับฟิลเตอร์แบบ LoG ที่หาได้จากวิธีดั้งเดิม ในที่นี้จะแสดงเฉพาะด้านบนซ้ายของฟิลเตอร์ด้วยขนาด 1 ใน 4 เท่านั้น เพราะมีลักษณะสมมาตรของแกนหมุน (rotational symmetric) ผลรวมของสัมประสิทธิ์ของฟิลเตอร์เป็น 0 ค่าคงที่สเกลไม่มีความสัมพันธ์โดยตรงกับขนาดของตัวปฏิบัติการ แต่อย่างไรก็ตามค่าคงที่สเกลค่าใหญ่มักจะถูกใช้ และจะไม่มีการต้ดขอบของภาพอย่างไม่มีเหตุผลของสัมประสิทธิ์ก็จะไม่เกิดขึ้น และทำให้ยากต่อการสร้างตัวปฏิบัติการที่สมดุลกัน

สาเหตุที่เกิดความแตกต่างที่น้อยมากๆ ระหว่างค่าสัมประสิทธิ์ของภาพที่ได้จากการสร้างขึ้นใหม่ (Reconstruction) กับภาพที่ได้จากการใช้ฟิลเตอร์ดั้งเดิม (original filter) เป็นผลอันเนื่องมาจากการควอนไตซ์ (quantization effect) ของการแทนที่สัมประสิทธิ์ของฟิลเตอร์แถว และฟิลเตอร์หลักที่เป็นจำนวนจริงนั่นเอง

ตัวอย่าง: ฟิลเตอร์แถว และฟิลเตอร์หลัก ==>

ฟิลเตอร์หลัก $h_1(\ ) = [\ -1\ -6\ -17\ -17\ 18\ 46\ 18\ -17\ -17\ -6\ -1\ ]$

ฟิลเตอร์แถว $h_2(\ ) = [\ 0\ 1\ 5\ 17\ 36\ 46\ 36\ 17\ 5\ 1\ 0\ ]$

จะได้ฟิลเตอร์แบบ LoG ที่สร้างขึ้นใหม่ (Reconstructed LoG filter) จากการใช้การแยกฟิลเตอร์แบบ LoG (Separate LoG filter)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ได้ดังนี้แสดงเฉพาะควอนแดนท์ด้านบนซ้าย (upper left quandant) ===>

0	-1	-5	-17	-36	-46
-1	-12	-47	-149	-198	-230
-5	-47	-170	-334	-522	-552
-17	-119	-374	-578	-306	0
-36	-198	-522	-306	1296	2484
-46	-230	-552	0	2484	4332

ถ้าใช้ฟิลเตอร์แบบ LoG ในแบบดั้งเดิม (Original LoG filter)
จะได้ดังนี้ แสดงเฉพาะควอนแดนท์ด้านบนซ้าย ===>

0	0	0	0	0	0	0	0	
0	0	0	0	-1	-2	-3	-4	
0	0	0	-1	-5	-17	-36	-46	==
0	0	-1	-12	-47	-149	-198	-230	==
0	-1	-5	-47	-170	-334	-522	-552	==
0	-2	-17	-119	-374	-578	-306	0	==
0	-3	-36	-198	-522	-306	1296	2484	==
0	-4	-46	-230	-552	0	2484	4332	==

==========

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.4 การทำนายทิศทางของภาพ (Predicate Based)

ในอัลกอริทึมนี้ใช้หาจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์ ใน 2 ทิศทางขนาดย่าน 3x3 ในภาพที่คำนวนแล้ว เพื่อหาว่าเข้ากับแบบใดใน 11 แบบ แบบที่พบจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์ เรียกว่า แบบ A-K ที่กำหนดตำแหน่งขอบของภาพได้ 24 ตำแหน่ง เครื่องหมายบวก และ/หรือ ลบ ในแผ่นจำลอง (template) แสดง ค่าบวก และ/หรือ ค่าลบ ต้องแสดงที่ตำแหน่งในภาพที่ผ่านการกรองแล้วเพื่อจับคู่กัน ตำแหน่งหน้าต่างที่ให้เป็น a-i จากซ้ายไปขวา บนไปล่าง ตัวอักษรแสดงว่าตำแหน่งที่ให้มาแสดงค่าที่ถูกทดสอบ *ตำแหน่ง* ของขอบของภาพ ในหน้าต่างถูกแสดงโดยพิเซลที่มีวงกลมและสัมพันธ์กับ *ทิศทาง* ที่แสดงโดยลูกศรที่อยู่ถัดไป

การทำนายประกอบด้วย 2 กลุ่มหลัก **กลุ่มแรก (ทำนายแบบ A-D)** ทำกับค่าศูนย์ในเอาท์พุทของการทำคอนโวลูชัน การทำนายแต่ละแบบมีทิศทางดังนี้ ทำนายแบบ A ทางใต้ แบบ B ทางเหนือ แบบ C ทางตะวันตก แบบ D ทางตะวันออก ส่วนทิศทางทแยง ต้องพิจารณาว่าเป็นแบบทั้งทางตั้ง และทางนอนหรือเปล่า การทำนาย 4 แบบนี้ A-D และ ครอบคลุมทั้ง 8 ทิศทางที่เป็นไปได้ของจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์ที่พิเซลตรงกลาง ตอนแรกเราตรวจสอบเครื่องหมายก่อนเพื่อหาแบบการทำนายที่เข้ากัน ถ้าไม่มีแบบที่เข้ากันได้แล้ว ก็จะไม่มีจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์ที่เหมาะสมกัน แต่ถ้าจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์ ถูกจับได้ที่พิเซลตรงกลาง ดังแสดงโดยพิเซลที่มีวงกลมในตำแหน่ง e กรณีย่อยของการทำนายที่เข้ากัน อันที่นำไปสู่การเพิ่มการตรวจสอบเครื่องหมายเท่าๆกับค่าของพิเซล นำไปใช้เพื่อการกำหนดทิศทางขอบของภาพ

กลุ่มที่ 2 ของการทำนายขอบ(E-K) คิดที่ค่าบวกเปลี่ยนไปเป็นค่าลบในเอาท์พุทของการทำคอนโวลูชัน เราเลือกกำหนดย่าน 3x3 ของพิเซลที่มีค่าบวกเพื่อหาการเปลี่ยนเป็นค่าลบ การทำนายที่เข้ากับแบบ E-K เราจับกับเครื่องหมายก่อนเหมือนกับกลุ่มที่ 1 โดยที่ไม่จำเป็นต้องไปพิจารณาค่าที่มุมทั้ง 4 ของย่าน3x3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทำนายแบบ A-D (Predicate A-D)มีค่าเอาท์พุท = 0 ที่ตำแหน่ง e

การทำนายแบบ A (Predicate A) : ตำแหน่ง d มีค่าเป็นบวก และ f มีค่าเป็นลบ โดยไม่คำนึงถึงค่าอื่นในย่าน 3x3 ซึ่งใช้สำหรับอธิบายทิศทางของขอบของภาพโดยมีทิศทางหลักเป็นทางใต้ แต่ทิศทางของขอบของภาพที่ดีกว่า อาจจะทำโดยพิจารณาค่าที่พิเซล b กับ h ด้วย แล้วรูปทรงนี้ก็จะเข้ากับการทำนายแบบ Aได้

	b	
+	0	-
	h	

<=== การทำนายแบบ A

กรณี 1 ===>

b ==> +

h ==> -

	+	
+	0	-
	-	

จะได้ทิศทางตะวันตกเฉียงใต้ และจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ e

==> กรณี 2

b ==> -

h ==> +

	-	
+	0	-
	+	

กรณี 3 ==>

b ==> -

h ==> -

	-	
+	0	-
	-	

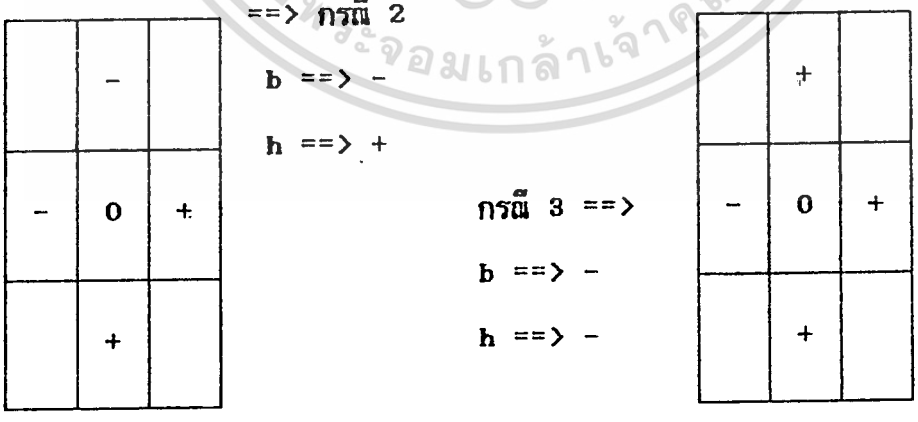
กรณี 2 จะได้ทิศทางตะวันออกเฉียงใต้ และจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ e

กรณี 3 จะได้ทิศทางใต้ และจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ e

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทำนายแบบ B (Predicate B) : ตำแหน่ง d มีค่าเป็นลบ และ f มีค่าเป็นบวก ใช้สำหรับอธิบายทิศทางของขอบของภาพโดยมีทิศทางหลักเป็นทางเหนือ แต่ทิศทางของขอบของภาพที่ดีกว่า อาจจะทำโดยพิจารณาค่าที่พิเซล b กับ h ด้วยแล้วรูปทรงนี้ก็จะเข้ากับการทำนายแบบ B

กรณี 1 จะได้ทิศทางตะวันตกเฉียงเหนือ และจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ e

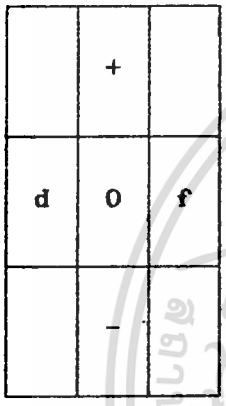


กรณี 2 จะได้ทิศทางตะวันออกเฉียงเหนือ และจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ e

กรณี 3 จะได้ทิศทางเหนือ และจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ e

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทำนายแบบ C (Predicate C) : ตำแหน่ง b มีค่าเป็นบวก และ h มีค่าเป็นลบ ใช้สำหรับอธิบายทิศทางของขอบของภาพ โดยมีทิศทางหลักเป็นทิศตะวันตก แต่ทิศทางของขอบของภาพที่ดีกว่าอาจจะทำโดยพิจารณาค่าที่พิเซล d กับ f ด้วย แล้วรูปทรงนี้ก็จะเข้ากับการทำนายแบบ C ได้

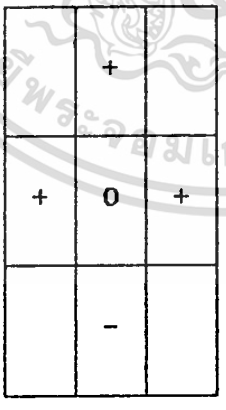


<=== การทำนายแบบ C

กรณี 1 ===>

d ==> +

f ==> +



กรณี 1 จะได้ทิศทางตะวันตกและจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ e

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กรณี 2 ===>

d ==> -

f ==> -

	+	
-	0	-
	-	

กรณี 2 จะได้ทิศทางตะวันตกและจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ e

การทำนายแบบ D (Predicate D) : ตำแหน่ง b มีค่าเป็นลบ และ h มีค่าเป็นบวก ใช้สำหรับอธิบายทิศทางของขอบของภาพ โดยมีทิศทางหลักเป็นทิศตะวันออก แต่ทิศทางของขอบของภาพที่ดีกว่าอาจจะทำโดยพิจารณาค่าที่พิเซล d กับ f ด้วยแล้ว รูปทรงนี้ก็จะเข้ากับการทำนายแบบ D

	-	
d	0	f
	+	

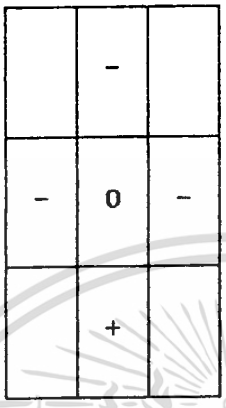
<=== การทำนายแบบ D

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กรณี 1 ===>

d ==> -

f ==> -

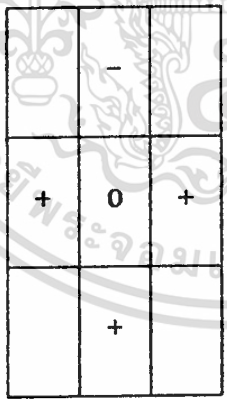


กรณี 1 จะได้ทิศทางตะวันออกและจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ e

กรณี 2 ===>

d ==> +

f ==> +



กรณี 2 จะได้ทิศทางตะวันออกและจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ e

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทำนายแบบ E (Predicate E) : ตำแหน่ง d มีค่าเป็นลบ และe,f มีค่าเป็นบวก ใช้สำหรับอธิบายทิศทางของขอบของภาพโดยสามารถแยกได้เป็น 3 กรณีย่อย ได้ดังนี้

	b	
-	+	+
	h	

<=== การทำนายแบบ E

	+	
-	+	+
	-	

<=== กรณีที่ 1

	+	
-	e	+
	h	

|h| < e ====>

จะได้ทิศทางตะวันตกเฉียงเหนือและจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง h

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$|d| < e$ ====>

	+	
d	+	+
	−	

จะได้ทิศทางตะวันตกเฉียงเหนือและจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง d

ถ้าไม่ใช่2แบบดังกล่าว

จะได้ ====>

	+	
d	+	+
	−	

จะได้ทิศทางตะวันตกเฉียงเหนือและจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง e

	−	
−	+	+
	+	

<=== กรณีที่ 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$|b| < e$ ====>

	b	
−	e	+
	+	

จะได้ทิศทางตะวันออกเฉียงเหนือและจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง b

$|d| < e$ ====>

	−	
d	e	+
	+	

จะได้ทิศทางตะวันออกเฉียงเหนือและจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง d

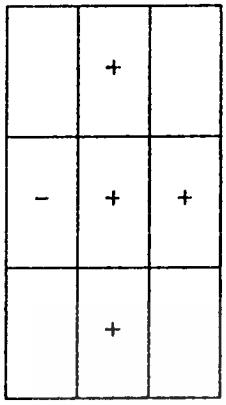
ถ้าไม่ใช่2แบบดังกล่าว

จะได้ ====>

	−	
−	e	+
	+	

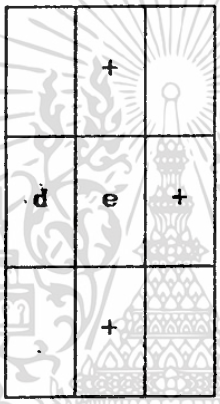
จะได้ทิศทางตะวันออกเฉียงเหนือและจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง e

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



<=== กรณีที่ 3

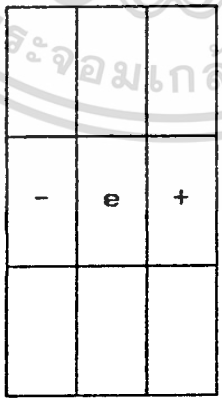
|d| < e ====>



จะได้ทิศทางเหนือและจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง d

ถ้าไม่ใช่แบบแรกแล้ว

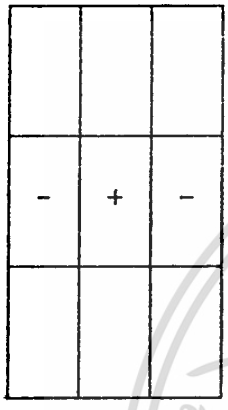
จะได้ ====>



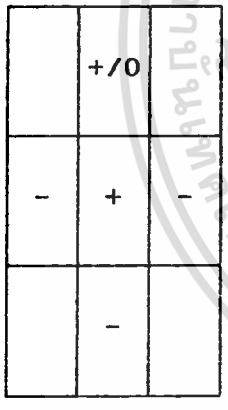
จะได้ทิศทางเหนือและจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง e

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

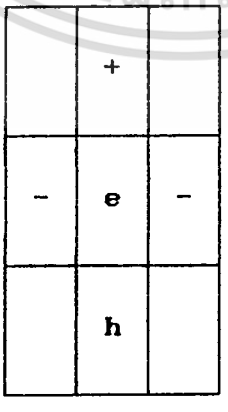
การทำนายแบบ F (Predicate F) : ตำแหน่ง d,f มีค่าเป็นลบ และ e มีค่าเป็นบวกใช้สำหรับอธิบายทิศทางของขอบของภาพโดยสามารถแยกได้เป็น 3 กรณีย่อยได้ดังนี้



<=== การทำนายแบบ F



<=== กรณีที่ 1



|h| < e ====>

จะได้ทิศทางตะวันตกเฉียงเหนือและจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง h

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ถ้าไม่ใช่แบบแรกแล้ว

จะได้ ====>

	+	
-	+	f

และ ====>

	+	
-	e	-

จะได้ทิศทางตะวันตกและจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง e และตำแหน่ง f

	-	
-	+	-
	+/0	

<=== กรณีที่ 2

====>

	-	
-	e	-
	+/0	

และ ====>

	-	
-	+	f
	+/0	

จะได้ทิศทางตะวันออกและจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง e และตำแหน่ง f

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$|d| < e$ ====>

	-	
d	+	-

จะได้ทิศทางเหนือและจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง d

กรณีที่ 3 ====>

ถ้าไม่ใช่ 2 กรณีแรก

-	+	f

จะได้ทิศทางใต้และจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง f

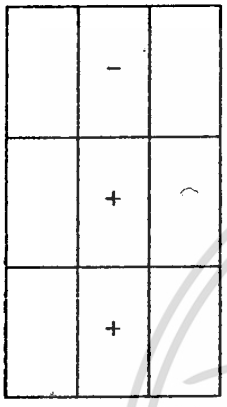
และ ====>

-	e	-

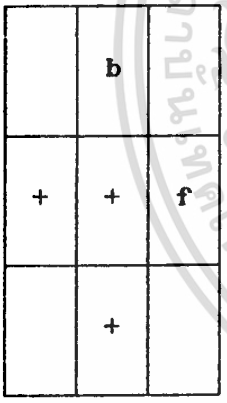
จะได้ทิศทางเหนือและจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง e

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทำนายแบบ G (Predicate G) : ตำแหน่ง b มีค่าเป็นลบ และe,h มีค่าเป็นบวก ใช้สำหรับอธิบายทิศทางของขอบของภาพโดยสามารถแยกได้เป็น 3 กรณีย่อย ได้ดังนี้



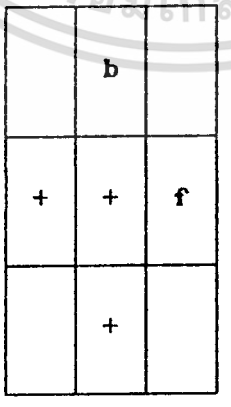
<=== การทำนายแบบ G



<=== กรณีที่ 1

$|b| < e$

ถ้า f > 0 ====>



จะได้ทิศทางตะวันออกและจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง b

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ถ้า f < 0 ====>

	b	
+	+	f
	+	

จะได้ทิศทางตะวันออกเฉียงใต้และจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง b

	–	
+	+	f
	+	

<=== กรณีที่ 2

f < 0, |f| < e ===>

	–	
+	+	f
	+	

จะได้ทิศทางตะวันออกเฉียงใต้และจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง f

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$f < 0, |f| > e$ ===>

	b	
+	+	f
	+	

จะได้ทิศทางตะวันออกเฉียงใต้และจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง e

	–	
	e	
	+	

<=== กรณีที่ 3

ถ้าไม่ใช่ 2 กรณีแรก

จะได้ทิศทางตะวันออกและจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง e

การทำนายแบบ H (Predicate H): ตำแหน่ง b,h มีค่าเป็นลบ และ e มีค่าเป็นบวก ใช้สำหรับอธิบายทิศทางของขอบของภาพโดยสามารถแยกได้เป็น 3 กรณีย่อย ได้ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

	-	
	+	
	-	

<=== การทำนายแบบ H

	-	
	+	f
g	h	

<=== กรณีที่ 1

f < 0

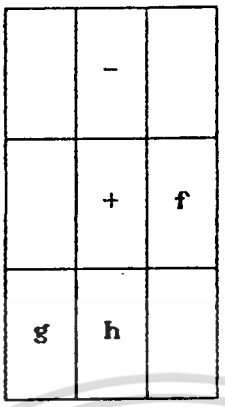
ถ้า g = 0 ====>

	-	
	+	f
g	h	

จะได้ทิศทางตะวันตกเฉียงใต้และจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง h

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

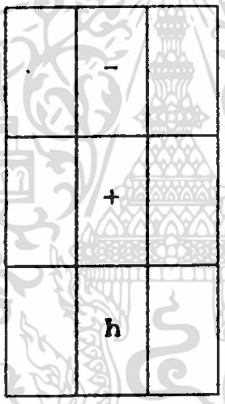
ถ้า g > 0 ====>



จะได้ทิศทางใต้และจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง h

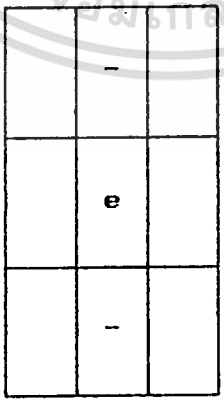
กรณีที่ 2

ถ้า g < 0 ====>



จะได้ทิศทางตะวันออกและจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง h

และ ====>



จะได้ทิศทางใต้และจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง e

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กรณีที่ 3

ถ้า f > 0 ====>

	-	
	+	
	h	

จะได้ทิศทางตะวันตกและจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง h

การทำนายแบบ I (Predicate I) : ตำแหน่ง f มีค่าเป็นลบ และe,h มีค่าเป็นบวก ใช้สำหรับอธิบายทิศทางของขอบของภาพโดยสามารถแยกได้เป็น 2 กรณีย่อย ได้ดังนี้

	+	-
	+	

<=== การทำนายแบบ I

กรณีที่ 1

ถ้า |f| < e ===>

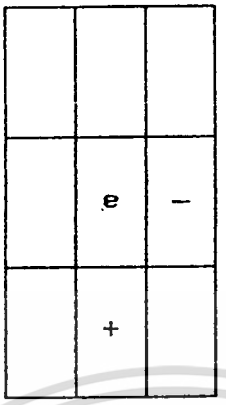
	e	f
	+	

จะได้ทิศทางใต้และจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง f

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

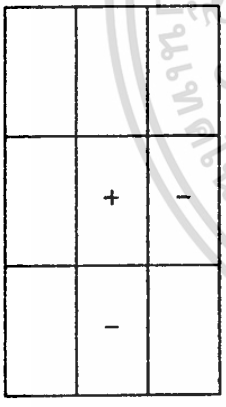
กรณีที่ 2

ถ้าไม่ใช่กรณีที่ 1 ===>



จะได้ทิศทางใต้และจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง e

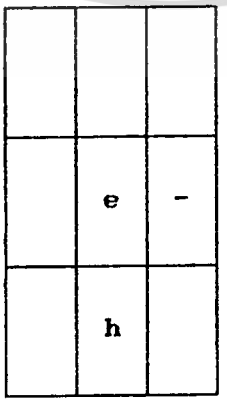
การทำนายแบบ J (Predicate J) : ตำแหน่ง f,h มีค่าเป็นลบ และe มีค่าเป็นบวก ใช้สำหรับอธิบายทิศทางของขอบของภาพโดยสามารถแยกได้เป็น 4 กรณีย่อย ได้ดังนี้



<=== การทำนายแบบ I

กรณีที่ 1

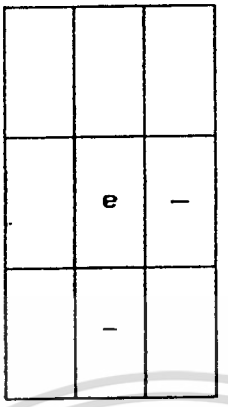
ถ้า |h| < e ===>



จะได้ทิศทางตะวันตกและจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง h

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

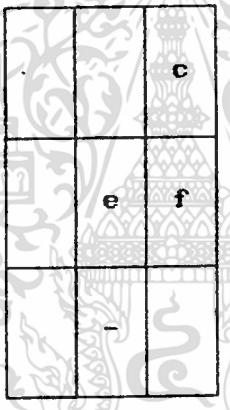
ถ้าไม่ใช่กรณีที่ 1 ===>



จะได้ทิศทางตะวันตกเฉียงใต้และจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง e

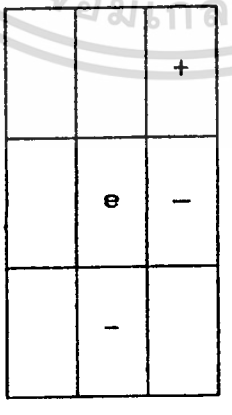
กรณีที่ 2

ถ้า c > 0, $|f| < e$ ===>



จะได้ทิศทางตะวันตกเฉียงใต้และจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง f

ถ้าไม่ใช่กรณีที่ 2 ===>

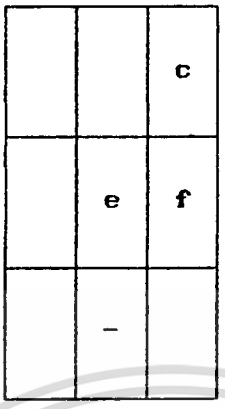


จะได้ทิศทางตะวันตกเฉียงใต้และจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง e

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

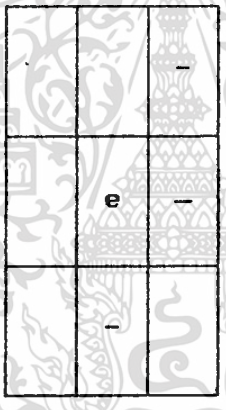
กรณีที่ 3

ถ้า c < 0, $|f|$ < e ===>



จะได้ทิศทางใต้และจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง f

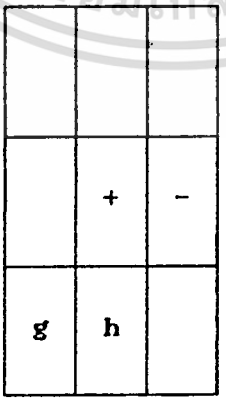
ถ้าไม่ใช่กรณีที่ 3 ===>



จะได้ทิศทางใต้และจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง e

กรณีที่ 4

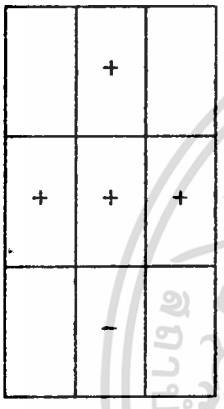
ถ้า g > 0 ===>



จะได้ทิศทางตะวันตกเฉียงใต้และจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง h

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

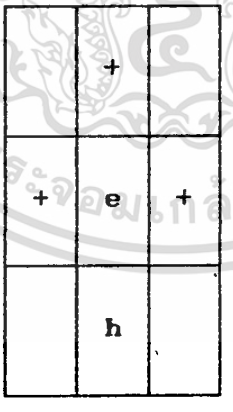
การทำนายแบบ K (Predicate K) : ตำแหน่ง h มีค่าเป็นลบ และ b,d,e,f มีค่าเป็นบวก ใช้สำหรับอธิบายทิศทางของขอบของภาพโดยสามารถแยกได้เป็น 2 กรณีย่อยได้ดังนี้



<=== การทำนายแบบ K

กรณีที่ 1

ถ้า |h| < e ===>

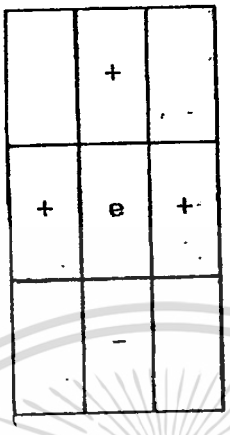


จะได้ทิศทางตะวันตกและจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง h

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กรณีที่ 2

ถ้าไม่ใช่กรณีที่ 1 ===>



จะได้ทิศทางตะวันตกและจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์อยู่ที่ตำแหน่ง e

3.5 การทดลองและผลการทดลอง

จากในภาคทฤษฎี เรานำมาเขียนเป็นโฟลว์ชาร์ต และโปรแกรมซึ่งเราใช้โปรแกรมภาษาซีในการพัฒนา โดยเขียนตามอัลกอริทึมดังนี้

อัลกอริทึมของโปรแกรม

1. จองหน่วยความจำขนาด 128 กิโลไบต์

2. รับข้อมูลจากภายนอก ขนาด 256x256 Bytes

3. หาฟิลเตอร์ โดยแยกเป็น 2 มิติ ตามทฤษฎี ขนาด 11x11

4. นำภาพอินพุทมาทำคอนโวลูชันในแนวแกน x กับ ฟิลเตอร์แบบแถว (Row LoG Filter)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.นำผลลัพธ์ที่ได้จาก (4) มาทำคอนโวลูชันในแนวแกน y กับฟิลเตอร์แบบคอลัมภ์ (Column LoG Filter)

6.ตรวจสอบค่าที่มีการเปลี่ยนแปลงผ่านค่าศูนย์ โดยใช้การทำนายทิศทางจากหัวข้อ 3.4

7.แสดงผลที่ได้ และนำไปประยุกต์ใช้ตามต้องการ

จากการทดลองที่แสดงจะพบว่า รูปโอยา (hoya) รูปที่ 5.3 แสดงภาพอินพุท และนำมาทำคอนโวลูชันด้วยค่า เบี่ยงเบนเท่ากับ 0.146 ค่าคงที่สเกลเท่ากับ 350 รูปที่ 5.4 และค่าเบี่ยงเบนเท่ากับ 2 ค่าคงที่สเกลเท่ากับ 350 ในรูปที่ 5.5

ถ้า ค่าเบี่ยงเบนเท่ากับ 0.146 ค่าคงที่สเกลเท่ากับ 350 จะได้ฟิลเตอร์แบบแถวและฟิลเตอร์แบบคอลัมภ์

h1 [ ] = [0 0 0 0 0 19 0 0 0 0 0]

h2 [ ] = [0 0 0 0 0 19 0 0 0 0 0]

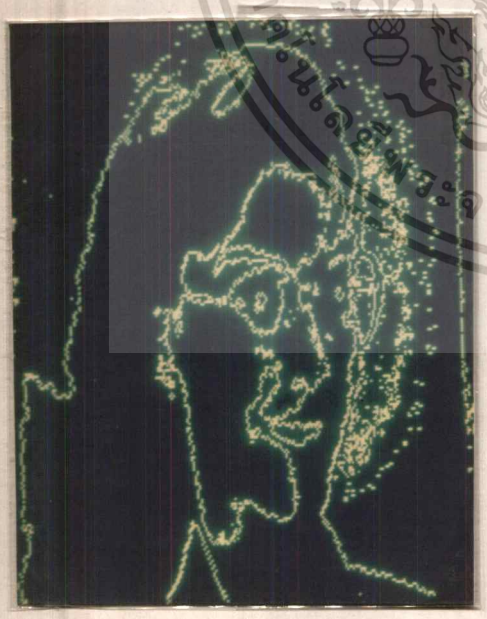
ถ้า ค่าเบี่ยงเบนเท่ากับ 2 ค่าคงที่สเกลเท่ากับ 350 จะได้ฟิลเตอร์แบบแถวและฟิลเตอร์แบบคอลัมภ์

h1 [ ] = [-4 -6 -6 0 10 16 10 0 -6 -6 -4]

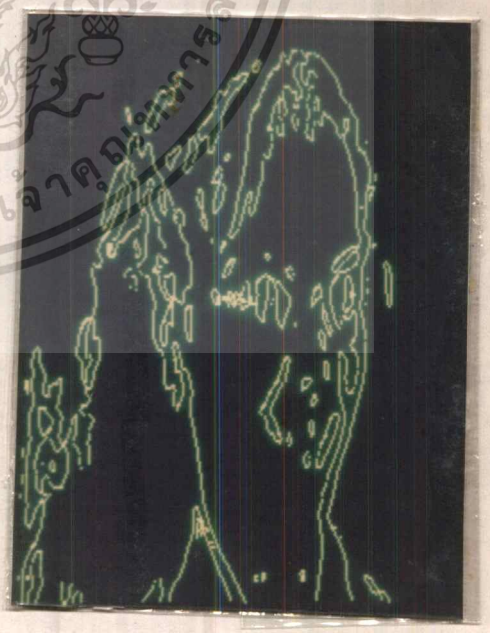
h2 [ ] = [1 2 5 10 14 16 14 10 5 2 1]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 3.3 ภาพโอยา ซึ่งเป็นภาพอินพุท



รูปที่ 3.4



รูปที่ 3.5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จะเห็นว่า เมื่อค่าเบี่ยงเบนเท่ากับ 2 ค่าคงที่สเกลเท่ากับ 350 ฟิลเตอร์จะมีลักษณะการกระจายเป็นแบบเกาส์เซียน ซึ่งเป็นลักษณะของตัวปฏิบัติการที่ดีในการตรวจจับขอบของภาพ ส่วนเมื่อใช้ค่าเบี่ยงเบนเท่ากับ 0.146 ค่าคงที่สเกลเท่ากับ 350 จะได้ฟิลเตอร์ที่มีลักษณะการหาขอบจากจุดที่ค่าความเข้มสูงสุด แต่การตรวจจับขอบที่ดีจะต้องตรวจจับจากจุดที่มีการเปลี่ยนแปลงความเข้มของเครื่องหมาย บวก เป็น ลบ หรือ ลบ เป็น บวก ดังนั้นจะพบว่ารูป 5.5 จะเป็นภาพที่แสดงขอบของภาพได้ดีกว่ารูป 5.4 และจากการทดลองหลาย ๆ ครั้ง หลาย ๆ รูป จึงสามารถสรุปได้ว่า ค่าเบี่ยงเบนที่ใช้ควรมีค่าน้อย แต่ไม่น้อยจนกระทั่งทำให้พารามิเตอร์ (parameter) ของมาส์ก LoG ไม่เป็นการกระจายแบบเกาส์เซียน ดังเช่น พารามิเตอร์ ของรูป 5.4 เช่นกันถ้าเราใช้ค่าเบี่ยงเบนที่มีขนาดมากเกินไป ก็มีผลต่อการกระจายของเกาส์เซียน ทำให้ลดรายละเอียดของภาพมากเกินไป ส่วนค่าคงที่สเกลที่ใหญ่เกินไป ก็จะเก็บรายละเอียดของภาพได้ดีกว่าค่าค่าคงที่สเกลขนาดเล็ก ดังนั้น การตรวจจับขอบให้ได้ภาพที่ดี จะต้องพิจารณาการกระจายแบบเกาส์เซียนของฟิลเตอร์แบบแถว และแบบคอลัมภ์ให้มีลักษณะดีที่สุด แต่ทั้งค่าคงที่สเกล และค่าเบี่ยงเบนที่ใช้ในมาส์ก LoG ในสมการที่ 2 จะมีช่วงการใช้งานได้กว้างกว่าการใช้ค่าเบี่ยงเบนเพียงค่าเดียวในมาส์ก LoG ในสมการที่ 1

3.6 สรุป

การคำนวนหาจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์ ด้วยวิธีของ LoG ต้องทำตามขั้นตอนดังนี้

1. ทำการคอนโวลูชัน ภาพ I ขนาด RxC ด้วย ฟิลเตอร์แบบ LoG ขนาด M^2 แล้วจะได้ภาพที่ถูกฟิลเตอร์ เป็น I_{LoG} ขนาด RxC

2. หาภาพที่มีจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์ ให้ชื่อเป็น I_{zero} ขนาด RxC โดยการหาตำแหน่งของจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์ใน I_{LoG} โดยใช้อัลกอริทึมดังกล่าว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.จากค่าของจุดเปลี่ยนแปลงที่ผ่านค่าศูนย์ แต่ละตัวและย่านรอบๆอีก 8 ตำแหน่ง นำมาคำนวณหาทิศทาง โดยใช้การทำนายขอบของภาพที่แสดงในหัวข้อ 3.4

จากการทดลองพบว่า ถ้าเราใช้ฟิลเตอร์ในสมการที่ 3.1 ซึ่งมีเฉพาะค่าเบี่ยงเบนเป็นพารามิเตอร์ จะทำให้ภาพที่ได้ สามารถตรวจจับขอบได้ในการเปลี่ยนแปลงช่วงค่าเบี่ยงเบนเล็กๆ เท่านั้น เพราะว่าถ้ามีการเปลี่ยนแปลงค่าเบี่ยงเบนเพียงนิดเดียวก็อาจจะทำให้ภาพที่ได้หาขอบของภาพไม่พบ ค่าเบี่ยงเบนนั้นจะมีผลต่อความไวของภาพมาก แต่ถ้าใช้ฟิลเตอร์แบบ LoG แล้วเราจะสามารถกำหนดค่าเบี่ยงเบนได้ในช่วงที่กว้างมาก เพราะว่า มีค่าคงที่สเกลมาช่วยลดความไวของฟิลเตอร์ในสมการ 3.1 เนื่องจากค่าคงที่สเกล และค่าเบี่ยงเบนเป็นอิสระต่อกัน จึงทำให้การเปลี่ยนแปลงของค่าเบี่ยงเบนไม่มีผลต่อภาพมากนัก เพราะฉะนั้นสรุปได้ว่า การเปลี่ยนแปลงค่าเบี่ยงเบนให้ผลต่อความไวของภาพ ส่วนการเปลี่ยนแปลงค่าคงที่สเกลจะมีผลในการลดความไวของภาพให้แสดงภาพได้ดีขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

ระบบการจดจำลายนิ้วมือ

Fingerprint Recognition System

ระบบจดจำลายนิ้วมือเป็นระบบที่จะตรวจสอบ และพิสูจน์ลายนิ้วมืออัตโนมัติ ระบบจดจำลายนิ้วมือในบทนี้จะกล่าวถึง พื้นฐานเบื้องต้นของการพิสูจน์ลายนิ้วมือ ความจำเป็นในการใช้ระบบจดจำลายนิ้วมือ จากนั้นส่วนต่อไปจะกล่าวถึงขั้นตอนต่างๆของระบบจดจำลายนิ้วมือ จากนั้นจะกล่าวถึงการทดลองและผลการทดลอง และสุดท้ายจะกล่าวถึงสรุปผลการทดลองและแนวทางการพัฒนาระบบให้ดียิ่งขึ้น

4.1 พื้นฐานเบื้องต้น [3]

ลักษณะที่เป็นเอกลักษณ์เฉพาะของแต่ละบุคคลอย่างหนึ่งคือลายนิ้วมือ เพราะลายนิ้วมือของแต่ละบุคคลจะไม่เหมือนกัน ที่เป็นเช่นนี้เพราะลายนิ้วมือไม่มีรูปแบบที่แน่นอนตายตัว ไม่มีกฎเกณฑ์ว่าลายนิ้วมือของคนหนึ่งคนใดจะเป็นเช่นใด ลายนิ้วมือไม่ขึ้นกับอะไรทั้งสิ้น ไม่ขึ้นกับกรรมพันธุ์ หญิงหรือชาย เชื้อชาติ รูปร่างของลายนิ้วมือจะไม่เปลี่ยนแปลง ลายนิ้วมือมีลักษณะเป็นเส้นนูนมากมายลากจากด้านหนึ่งไปอีกด้านของนิ้วมือ แม้ว่าลายนิ้วมือจะมีรูปร่างไม่แน่นอน แต่เราสามารถจำแนกลายนิ้วมือแต่ละลายที่พบได้เป็น 4 ลักษณะใหญ่ๆ คือ เส้นโค้ง ลูป ก้นหอย และหลายอย่างประกอบกัน เส้นโค้ง หมายถึงลักษณะของลายที่ลากจากด้านหนึ่งไปอีกด้านหนึ่งของนิ้วมือ ลูป หมายถึง ลักษณะที่ลายหนึ่งลายหรือมากกว่า ลากจากด้านหนึ่งแล้ววกกลับไปออกที่ด้านเดิม ก้นหอย หมายถึงรูปแบบของลายนิ้วมือที่วนเข้าหาจุดศูนย์กลาง นอกจากนี้ ยังมีบางลายที่มีลักษณะข้างต้นมากกว่าหนึ่งลักษณะประกอบกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เนื่องจากลายนิ้วมือมีลักษณะเป็นหนึ่งเดียว ไม่มีใครมีลายนิ้วมือเหมือนกับอีกคนได้ ลายพิมพ์ของลายนิ้วมือ จึงสามารถนำไปใช้ประโยชน์เป็นตัวแทนบุคคลได้เป็นอย่างดี ซึ่งสามารถนำเอาคุณสมบัตินี้ไปใช้ประโยชน์ในงานต่างๆ ดังเช่น งานทะเบียนประชาชนงานปราบอาชญากรรม สิ่งสำคัญของงานที่ใช้ลายนิ้วมือเป็นตัวแทนบุคคล คือ การพิสูจน์ลายนิ้วมือ การที่จะพิสูจน์ว่าลายนิ้วมือคู่หนึ่งเป็นของคนๆเดียวกันหรือไม่ จะใช้คนตรวจสอบดูว่าลายทั้งสองเหมือนกันหรือไม่ ซึ่งเป็นการสูญเสียทั้งแรงงานและเวลา จึงนำมาสู่การพัฒนาขบวนการวิเคราะห์ เพื่อพิสูจน์ลายนิ้วมืออัตโนมัติ ซึ่งเป็นเรื่องที่ทางวิทยาศาสตร์ และทางส่วนการปราบอาชญากรรมให้ความสนใจมาก การเพิ่มความสามารถ และความเร็วในการวิเคราะห์ลายนิ้วมือ ทำให้เกิดการพัฒนาระบบจดจำลายนิ้วมือขึ้น เรียกว่า **ระบบการจดจำลายนิ้วมือ** (Fingerprint Recognition System) ระบบอัตโนมัตินี้ ต้องสามารถจะทำสิ่งต่างๆได้ดังนี้ คือ สามารถจับคู่ลายนิ้วมือที่เหมือนกันได้ และสามารถระบุได้ว่าลายนิ้วมือสองลาย เป็นลายนิ้วมือของคนๆเดียวกันหรือไม่ นอกจากนี้ ระบบอัตโนมัตินี้ ต้องคำนึงถึงเวลาในการค้นหา ซึ่งเป็นสิ่งที่สำคัญที่ต้องพิจารณา ยิ่งสามารถลดเวลาลงได้ยิ่งดีเท่านั้น

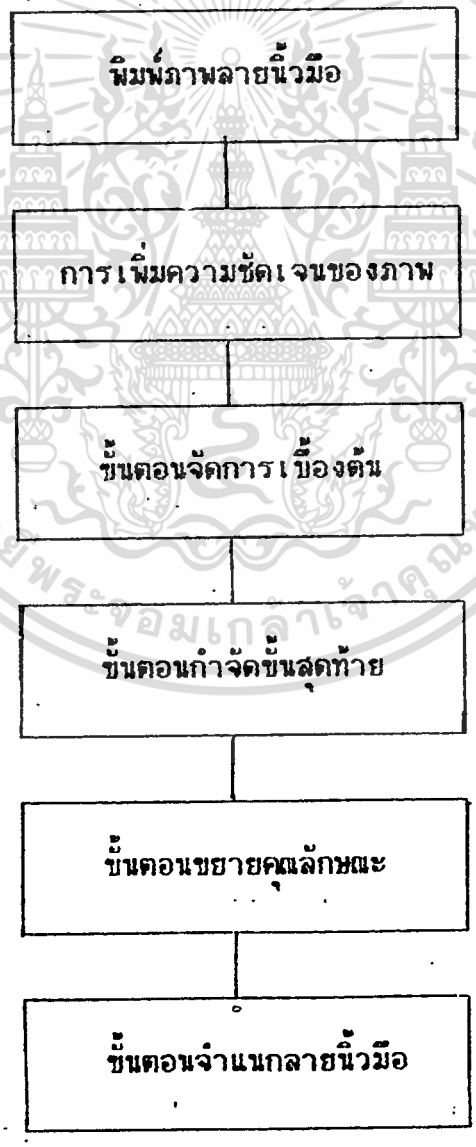
ระบบการจดจำลายนิ้วมือ ต้องสามารถเก็บข้อมูลจากลายนิ้วมือที่ไม่มีอะไรเป็นกฎเกณฑ์ จึงต้องหาข้อมูลที่แต่ละลายนิ้วมือมีไม่เหมือนกัน นั่นก็คือลักษณะโดดเด่นของแต่ละลายนิ้วมือ เนื่องจากเส้นลายนิ้วมือส่วนใหญ่ จะเป็นเส้นเดี่ยว และเริ่มที่ด้านข้างของนิ้วมือลากผ่านไป ออกที่ปลายอีกด้านของนิ้วมือ มีลักษณะต่างๆตามที่กล่าวไว้ข้างต้น แต่มีบางลายเส้นที่ไม่ได้เริ่มต้นหรือสิ้นสุด ที่ข้างใดข้างหนึ่งของนิ้วมือ หรือเป็นเส้นสั้นๆ หรือไม่ได้เป็นเส้นเดี่ยวตลอดลาย กล่าวคือเป็น เส้นที่เส้นเดี่ยวแล้วแตกออกเป็น 2 ลาย เรียกว่า เส้นกิ่ง (Branching) หรือเป็นเส้นเดี่ยว 2 เส้น รวมเป็นเส้นเดียว เรียกว่า เส้นที่รวมกัน (Merging) ลักษณะต่างๆเช่นนี้ เรียกว่า ลักษณะโดดเด่นของลายนิ้วมือ (Minutiae) ลายนิ้วมือหนึ่งๆ จะมีลักษณะเช่นนี้ ปรากฏอยู่ประมาณ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

50-100 แห่ง ระบบจดจำลายนิ้วมือ จึงพิจารณาไปที่ลักษณะโดดเด่นนี้ สัมพันธ์กับตำแหน่งที่เกิดลักษณะนั้น

4.2 ระบบจดจำลายนิ้วมือ [3]

ระบบจดจำลายนิ้วมือ สามารถแสดง ได้ตามแผนภาพต่อไปนี้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนแรก คือ **ขั้นตอนการเปลี่ยนภาพลายนิ้วมือมาเป็นภาพในแบบดิจิตอล** ขนาด 256x256 พิเซล ความสว่างขนาด 0-255 เราต้องพิมพ์ลายนิ้วมือก่อนแล้วนำลายพิมพ์ที่ได้มาเปลี่ยนเป็นภาพดิจิตอล เนื่องจากเราต้องการข้อมูล ที่ถูกต้องมากที่สุด พิมพ์ลายนิ้วมือ จึงควรที่จะชัดเจนสมบูรณ์ ข้อผิดพลาดน้อยที่สุด เพราะการเกิดข้อผิดพลาดขึ้น ก็คือ เราได้ข้อมูลผิด อาจนำไปสู่การพลาดการยอมรับลายนิ้วมือของคนๆเดียวกันได้ ข้อผิดพลาดที่สำคัญ ได้แก่ ลายขนานกันที่ต่อติดกัน เนื่องจาก หมึกพิมพ์เลอะ ลายไม่ต่อติดกัน เนื่องจาก หมึกพิมพ์ไม่ติด

ขั้นตอนต่อมา คือ **ขั้นตอนการเพิ่มความชัดเจนของภาพ** (Picture Enhancement) เราสามารถเพิ่มความชัดเจนของภาพได้ โดยการลดระดับความสว่างของแต่ละพิเซลให้เหลือเพียง 2 ระดับ เรียกว่า ภาพไบนารี (Binary Image) ภาพจะมีความชัดเจนขึ้น

ขั้นตอนที่สาม คือ **ขั้นตอนจัดการเบื้องต้น** (Preprocessing) ในขั้นตอนนี้ จะแบ่งภาพลายนิ้วมือ ออกเป็นภาพย่อยๆ (Pattern decomposition) เพื่อลดข้อมูลลง และเพื่อความสะดวก จากนั้นทำการเปลี่ยนลายนิ้วมือแต่ละลาย เป็นข้อมูลที่มีความสัมพันธ์กับทิศทางของลาย ซึ่งเป็นการนำมาซึ่งข้อมูลที่จะนำไปพิสูจน์ต่อไป

ขั้นที่สี่ คือ **ขั้นตอนกำจัดขั้นสุดท้าย** (Postprocessing) จากที่กล่าวไว้ข้างต้น การผิดพลาดจากการพิมพ์ลายนิ้วมือ เป็นสิ่งที่ให้ความสำคัญ ต้องพยายามให้มีน้อยที่สุด แต่เราไม่สามารถพิถีพิถัน กับการพิมพ์ลายนิ้วมือมากเกินไป เนื่องจากเป็นงานที่นำใช้กับคนหมู่มาก หรือลายนิ้วมือที่ได้มาจาก งานอาชญากรรม ซึ่งต้องนำเอามาใช้พิสูจน์ จึงกล่าวได้ว่า การพิมพ์ลายนิ้วมือมีขีดจำกัด เราจึงต้องมีขบวนการที่จะวิเคราะห์ และแก้ไขข้อผิดพลาดที่เกิดจากการพิมพ์ลายนิ้วมือขึ้น นั่นคือ ขั้นตอนที่ห้า

ขั้นตอนที่ห้า คือ **ขั้นตอนขยายคุณลักษณะ** (Feature Extraction) เมื่อผ่านขั้นตอนต่างๆข้างต้นมาแล้ว เราจะได้ข้อมูลที่ถูกต้องสมบูรณ์สามารถนำไปพิสูจน์ได้ ขั้นตอนนี้ ก็จะนำข้อมูลที่ได้มาพิสูจน์ความเป็นเฉพาะบุคคลของลายนิ้วมือ โดยพิจารณาที่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ลักษณะลายนิ้วมือและลักษณะโดดเด่น จะทำการแปลงข้อมูลที่ได้มาเป็นโค๊ตข้อมูลแสดงลักษณะลายนิ้วมือและลักษณะโดดเด่น

ขั้นตอนสุดท้าย คือ ขั้นตอนจำแนกลายนิ้วมือ (Classification) จะทำการแทนลักษณะโครงร่างลายนิ้วมือในแต่ละหน้าต่าง ด้วยข้อมูลที่แปลงมาจากขั้นตอนที่แล้ว แล้วทำการลดข้อมูลลง โดยการใช้ไวยกรณ์แบบทรี (Tree Gramma) จากนั้นก็ถึงส่วนการจดจำลายนิ้วมือ โดยการเก็บข้อมูลที่ได้ลงเป็นข้อมูลของบุคคลที่จดจำได้

4.2.1 การเพิ่มความชัดเจนของภาพ

ภาพลายนิ้วมือที่เราได้มาจากขั้นตอนแรก เป็นลายขนาด 256x256 ไบท์ ความสว่าง 0-255 ระดับ ตามการวิจัยโครงงานนี้ เราสามารถเพิ่มความชัดเจน โดยใช้วิธีการเฉลี่ยค่ารอบย่าน ดังที่กล่าวมาแล้วในบทที่ 2 หลังจากนั้น ทำการเปลี่ยนขนาดความสว่างให้เหลือ 2 ระดับ (Binary Pattern) โดยการตัดค่าเทรสโชลด์ ซึ่งใช้วิธีอีสโตรแกรม

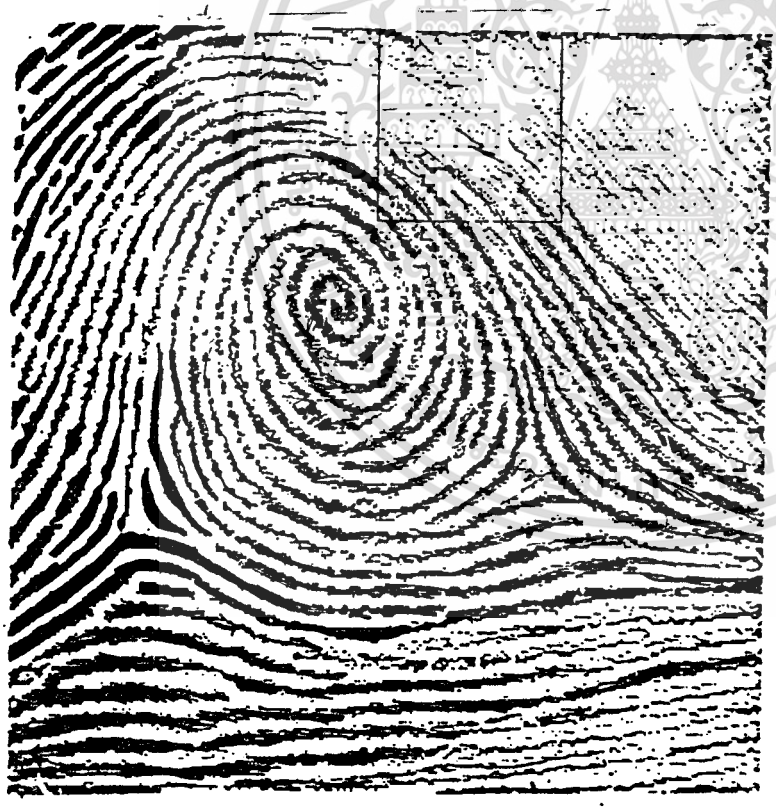
4.2.2 การแบ่งแยกรูปภาพ

ถึงแม้ว่า เราจะใช้วิธีอธิบายลายนิ้วมือ โดยการแทนลายโดยใช้ไวยากรณ์แบบทรี มันก็เป็นการยากที่จะทำโครงสร้างไวยากรณ์แบบทรี ออกมาแทนลายนิ้วมือได้ เนื่องจากลายนิ้วมือหนึ่งๆ จะมีลายอยู่ประมาณ 100 ลาย ซึ่งต้องการการเก็บข้อมูลจำนวนมาก และจะทำให้เกิดความสับสนได้ง่าย ทางหนึ่ง ที่จะแก้ปัญหาได้ ก็คือ การลดข้อมูลลง โดยการแบ่งลายนิ้วมือออกเป็นส่วนๆ ซึ่งสามารถทำโครงสร้างไวยากรณ์แบบทรีได้ง่ายกว่า โดยที่จะต้อง ให้สัมพันธ์กับส่วนที่มีลักษณะพิเศษกับส่วนที่ไม่มี อย่างไรก็ตาม ลายนิ้วมือไม่สามารถแยกส่วนออกมาได้ชัดเจน ดังนั้นการแยกส่วนโดยไม่มีกฎเกณฑ์อาจจะดีพอๆกับแบบอื่นได้ การแยกส่วนลายนิ้วมือจะต้องคำนึงถึงหลัก 2 ประการ คือ หนึ่ง -ขนาดเมื่อแยกออกมาต้องไม่เล็กเกินไป ทำให้มีส่วนที่แยกออกมามากเกินไป

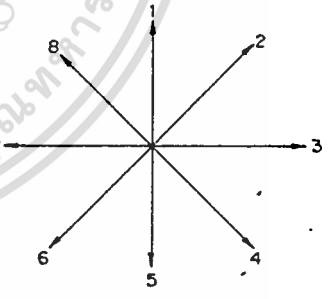
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สอง -ต้องไม่ใหญ่เกินไปจนมีลายในแต่ละส่วนมากเกินไป จนทำให้การแบ่งแยกไม่มีประโยชน์ เพราะข้อมูลในแต่ละส่วนมากเกินไป

ดังนั้น เราจึงแบ่งแยกลายนิ้วมือทั้งหมดออกมา 16 ส่วน เป็นเมตริกขนาด 4x4 หน้าต่าง โดยแต่ละส่วน เป็นหน้าต่างที่มีขนาด 48x48 พิเซล ดังในรูปที่ 4.1 การจำแนกลายนิ้วมือก็จะจำแนกแต่ละหน้าต่าง โดยไม่เกี่ยวกับหน้าต่างอื่น สิ่งที่ต้องคำนึงถึงต่อมา ก็คือ ตำแหน่งของหน้าต่างเป็นสำคัญ เพราะว่าการเลื่อนตำแหน่งของหน้าต่างอาจนำไปสู่การเคลื่อนย้ายข้อมูลจากหน้าต่างหนึ่งไปสู่อีกหน้าต่างได้ ซึ่งจะทำให้เกิดการพลาดในการจำแนกลายนิ้วมือในหน้าต่างนั้นได้



รูปที่ 4.1 แสดงลักษณะลายนิ้วมือ



รูปที่ 4.2 แสดงทิศทางของรหัสต่างๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

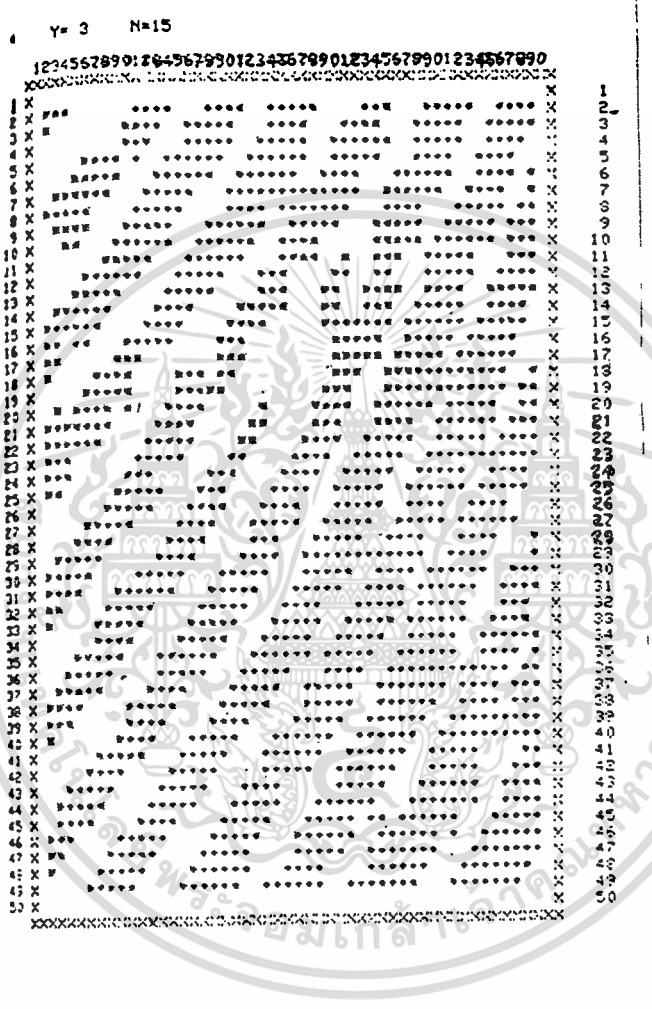
4.2.3 ขบวนการจัดการเบื้องต้น

จุดประสงค์ของขบวนการนี้ คือ การแทนภาพลายนิ้วมือมาเป็นข้อมูลที่นำไปใช้ได้ ลายแต่ละลายจะถูกวิเคราะห์แล้วเปลี่ยนเป็นข้อมูล ฉะนั้น เราต้องมีขบวนการที่สามารถสแกนลายนิ้วมือแต่ละเส้นได้ และข้อมูลที่เก็บมาได้นี้ ต้องเป็นข้อมูลที่มีประโยชน์และสามารถเอาไปใช้ได้ จากที่กล่าวข้างต้นการพิสูจน์ และจำแนกลายนิ้วมือจะพิจารณาที่ลักษณะโดดเด่นของลายนิ้วมือ ฉะนั้น ข้อมูลที่เก็บมาได้ ต้องประกอบด้วยข้อมูลที่แสดงลักษณะโดดเด่น ข้อมูลที่เก็บมาต้องเป็นข้อมูลที่จำเป็น ไม่ควรให้ข้อมูล ที่มีมากเกินความจำเป็น ลายนิ้วมือแต่ละเส้น จะมีลักษณะเป็นเส้นหนา และความหนาของแต่ละลายไม่เท่ากันทุกเส้น แต่ความหนาเป็นข้อมูลที่เราไม่สนใจ ดังนั้น เราจึงลดข้อมูลให้น้อยลงได้โดยการแทนเส้นลายนิ้วมือ ด้วยเส้นๆเดียว นั่นคือ การกำจัดความแตกต่างของความหนาข้อมูลที่เราสนใจ นอกเหนือจากลักษณะโดดเด่น ก็คือ ทิศทาง กับตำแหน่งของลายนิ้วมือ ทิศทางของลายนิ้วมือ คือ ลักษณะของลายเส้นว่าลากไปในทิศใด มีความสัมพันธ์กับตำแหน่งของลายอย่างไร ดังนั้นพอสรุปได้ว่า ข้อมูลที่ต้องสำหรับขบวนการนี้ ก็คือ ข้อมูลแสดงที่ลักษณะโดดเด่นของลายนิ้วมือ และข้อมูลทิศทางของลายนิ้วมือที่แสดงความสัมพันธ์กับตำแหน่ง

เมื่อได้ขอบเขตของข้อมูลที่ต้องการแล้ว เราจะแทนข้อมูลเหล่านี้ด้วยสัญญลักษณ์ดังนี้ ทิศทางของลายสามารถแสดงได้โดยใช้ตัวเลขที่แสดงถึงทิศทั้ง 8 ทิศ แสดงได้ดังรูปที่ 4.2 ส่วนลักษณะโดดเด่นสามารถได้โดยสัญญลักษณ์ดังนี้

+	หมายถึง	จุดเริ่มต้นเส้นลาย
.	หมายถึง	จุดสุดสุดเส้นลาย
*	หมายถึง	ลายสองลายบรรจบกัน
B	หมายถึง	ลายเดี่ยวแยกเป็นสองลาย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



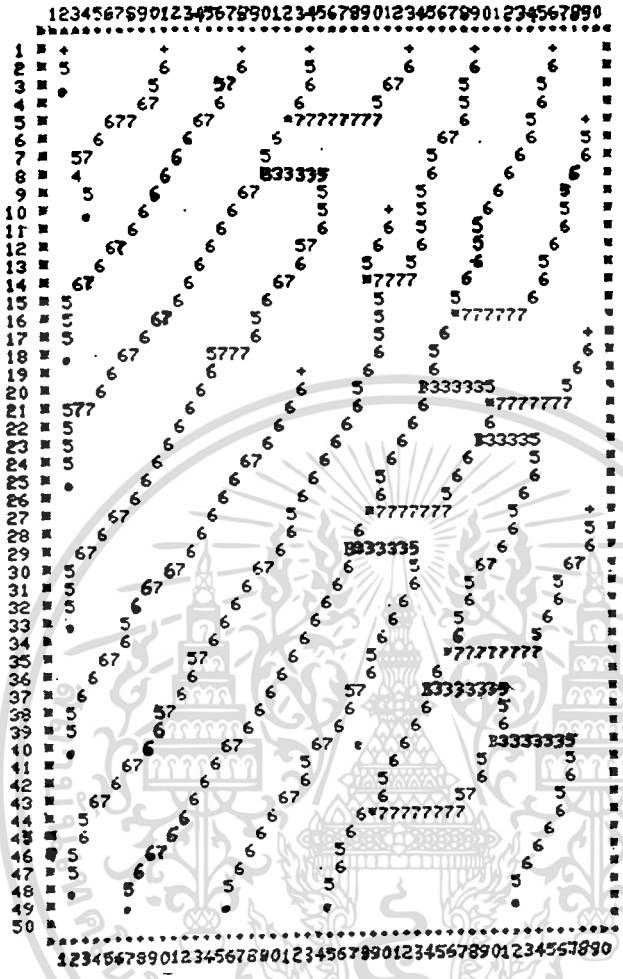
รูปที่ 4.3 แสดงหน้าต่างที่ 15 ของลายนิ้วมือของรูป 4.1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สัญญลักษณ์เหล่านี้ จะถูกแทนลงไป เมื่อพบข้อมูลนั้นในตำแหน่งที่เราสนใจ เพราะ เราต้องการข้อมูลที่สัมพันธ์กับตำแหน่ง ด้วยวิธีนี้ จะได้ตัวเลขมาแทนทิศทางลาย และสัญญลักษณ์มาแทนลักษณะโดดเด่นที่พบ ความหนาของลายถูกกำจัดให้หมดสิ้นไป ฉะนั้น ภาพลายนิ้วมือที่ผ่านขบวนการจัดการเบื้องต้นนี้ จะมีขนาดเท่าเดิม แต่จะเหลือแต่สัญญลักษณ์ที่แทนลงไปเท่านั้น ดังที่แสดงในรูปที่ 4.3

เทคนิคการให้ตัวเลขแสดงทิศทาง และสัญญลักษณ์แสดงลักษณะโดดเด่นเป็นสิ่งสำคัญที่ต้องพิจารณา ในขั้นตอน การสแกนลายนิ้วมือ จะเริ่มจากการตรวจในแนวนอนซ้ายไปขวา และในแนวตั้งบนไปล่าง เพื่อค้นหาลาย เมื่อพบก็จะให้สัญญลักษณ์ " + " แทนจุดเริ่มต้นลาย แล้วทำการไล่เก็บข้อมูลตลอดลายนั้น การสแกนลายนิ้วมือ จะไล่ไปตามขอบด้านซ้ายของลายนิ้วมือ แต่ละจุดที่ไล่ไปก็จะมีการให้ตัวเลขแสดงทิศทาง ตัวเลขทิศทางที่จุดหนึ่งๆ หมายถึง ทิศทางที่ลายลากไป การให้ตัวเลขทิศทางแบบนี้จะบอกว่าจุดต่อไปอยู่ในทิศทางใด อันเป็นประโยชน์ต่อการทำในขั้นตอนต่อไป เมื่อไม่สามารถให้ทิศทางต่อไปได้ก็หมายถึง สิ้นสุดลายแล้ว จะให้สัญญลักษณ์ " · " แทนจุดสิ้นสุด ในขณะที่ให้ทิศทางอยู่ก็ต้องมีการตรวจสอบเสมอว่า เกิดการแยกสาขาของลาย หรือเกิดลายๆอื่นมาบรรจบกับลายที่พิจารณาอยู่หรือไม่ ถ้ามีก็จะให้สัญญลักษณ์ " B " แทนในตำแหน่งลายที่แยกสาขาออกไป และสัญญลักษณ์ " * " ในตำแหน่งที่ลายอื่นมาบรรจบกับลายที่พิจารณา เมื่อให้ข้อมูลครบแล้ว ก็จะกลับไปสแกนเพื่อค้นหาลายอื่นๆ เมื่อพบก็จะไปไล่เก็บข้อมูลตลอดลายเหมือนกับที่กล่าวมาแล้วต่อไป ทำอย่างนี้จนสแกนสิ้นสุดทั้งหน้าต่าง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



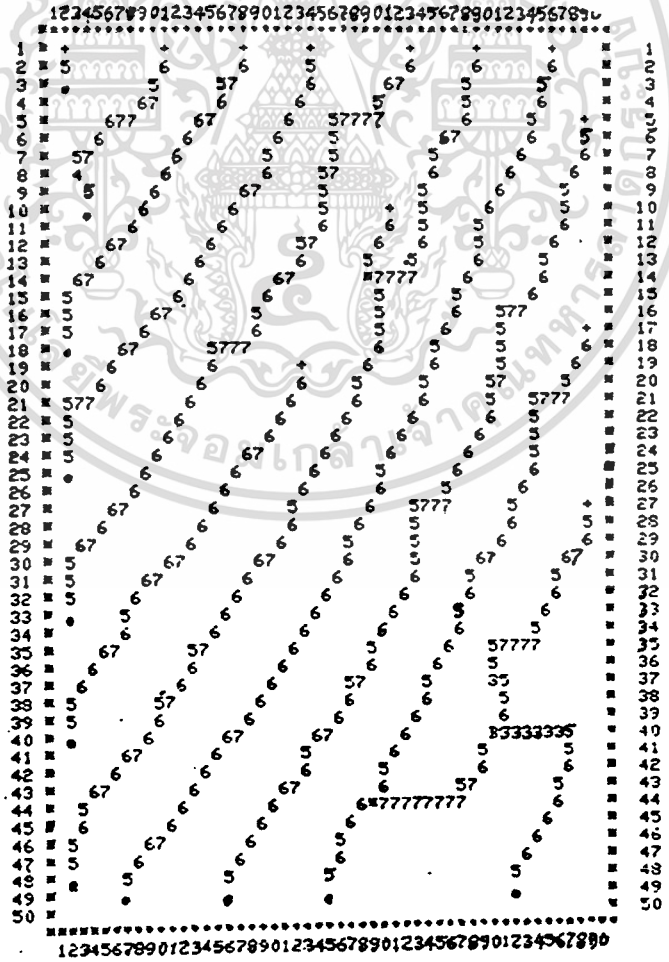
รูปที่ 4.4 แสดงรหัสหลังการจัดการเบื้องต้น

4.2.4 ขบวนการกำจัดขั้นสุดท้าย

ขบวนการส่วนนี้ เป็นส่วนที่สำคัญส่วนหนึ่ง ถ้าไม่มีส่วนนี้การพิสูจน์ลายนิ้วมือมีโอกาสที่ผิดพลาดได้มาก เนื่องจากการพิมพ์ลายนิ้วมือมีข้อจำกัด ข้อมูลที่ได้มา จึงไม่ใช่ข้อมูลที่ถูกต้องสมบูรณ์ จึงต้องมีขบวนการที่จะกำจัดข้อผิดพลาดที่เกิดจากการพิมพ์ลายนิ้วมือนั่นคือ ขบวนการกำจัดขั้นสุดท้าย จุดประสงค์ของขบวนการนี้ คือ ศึกษาข้อมูลที่ได้มาในแต่ละลาย แล้ววิเคราะห์เพื่อตัดสินใจว่า ลายสองลายที่เชื่อมติดกันเป็นการเชื่อมติดกันจริงๆ หรือเกิดจากการที่หมึกเลอะ ขบวนการกำจัดขั้นสุดท้าย ยังต้องสามารถลดช่องว่างเล็กๆ อันที่แบ่งลายหนึ่งลายออกเป็นสองลายหรือมากกว่าได้ อันเกิดจากบางส่วนของลายพิมพ์ไม่ติดได้ด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เทคนิคของขบวนการนี้ คือ การศึกษาที่ลักษณะโดดเด่นในแต่ละลาย การพิสูจน์ว่าลายสองลาย ที่เชื่อมกันควรจะเป็นเช่นนั้นจริงหรือไม่ จะทำโดยการดูที่ตำแหน่งของ " * " และ " B " ลายที่เกิดจากการเลอะของหมึก จะมีส่วนของลายที่เชื่อมกันจะเป็นส่วนสั้นๆ ฉะนั้น ตำแหน่งของ " * " และ " B " จะอยู่ใกล้กัน เมื่อพบลักษณะเช่นนี้จะสรุปว่า เป็นข้อผิดพลาดที่เกิดจากการเลอะของหมึก แล้วจะแยกจุดที่เชื่อมกันของลายทั้งสองออกจากกัน ในกรณีการพิสูจน์ว่าลายหนึ่งลายแยกเป็นหลายลายที่เกิดจากการพิมพ์ไม่ติดบางส่วน สามารถพิสูจน์ได้โดยการพิจารณา จุดสิ้นสุดกับจุดเริ่มต้นที่ใกล้ๆกัน แล้วเปรียบเทียบทิศทางโดยรวมของทั้งลายว่าใกล้เคียงกันหรือไม่ แล้วตัดสินใจว่าลายทั้งสองควรต่อกันหรือไม่ ดังที่แสดงในรูป 4.5



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 4.5 แสดงรหัสหลังการจัดการขั้นสุดท้าย

4.2.5 การขยายคุณลักษณะ

หลังจากที่ลายนิ้วมือผ่านขบวนการต่างๆมาแล้ว จะได้ข้อมูลของลายนิ้วมืออย่างครบถ้วนและสมบูรณ์ ที่สามารถนำไปพิสูจน์ได้ ตามที่กล่าวมาแล้วว่า ในลายนิ้วมือหนึ่งๆ จะมีลักษณะโดดเด่นอยู่ประมาณ 50-100 แห่งอยู่ในตำแหน่งต่างๆ ดังนั้นสามารถสร้างลายนิ้วมือที่แตกต่างกันได้ ทั้งหมด 2^{50} หรือเท่ากับ 10^{15} แบบ ดังนั้น ตัวแปลความหมายที่จะทำให้เราสามารถพิสูจน์ลายนิ้วมือได้ คือ เราต้องสามารถพิจารณาลักษณะโดดเด่นต่างๆ ที่ตำแหน่งต่างๆ ได้ เราจะสามารถพิสูจน์ลายนิ้วมือที่แตกต่างกันได้โดยไม่จำกัด และสามารถนำไปใช้ในทางปฏิบัติได้

ขั้นตอนนี้มีจุดประสงค์ คือ พิสูจน์ความเป็นเฉพาะบุคคลของลายนิ้วมือ ขั้นตอนนี้จะขยายขอบเขตของข้อมูลให้ได้มากขึ้น โดยการเปลี่ยนตัวเลขแสดงทิศทาง มาเป็นลายเส้นแสดงลักษณะของลายนิ้วมือ 4 แบบ A , B , C , D และใช้ไวยากรณ์แบบทรีเพื่ออธิบายลักษณะลายนิ้วมือในหน้าต่างนั้น

ข้อมูลลายนิ้วมือมีอยู่ 3 แบบที่สัมพันธ์กับโครงร่างของลายในแต่ละหน้าต่าง

1. ลายต่อเนื่อง ที่เริ่มจากขอบด้านหนึ่งของหน้าต่างไปออกอีกด้านของหน้าต่าง ลายลักษณะนี้ไม่มีลักษณะโดดเด่น และไม่มีข้อมูลให้พิสูจน์

2. ลายที่เริ่มต้นภายในหน้าต่าง ลายที่แตกสาขา ลายสองลายที่รวมเป็นลายเดียว ลายช่วงสั้นๆ ลายลักษณะนี้จะมีข้อมูลมากที่พิสูจน์

3. ลายที่อยู่ด้านข้างของลายอื่น แทนได้ด้วยสัญญลักษณ์ R หมายถึง ด้านขวา และ L หมายถึง ด้านซ้าย

จากที่อธิบายข้างต้นสามารถแสดงสัญญลักษณ์ได้ดังรูปที่ 4.6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ลักษณะลาย	โค๊ต
	A
	B
	C
	D
	X
	*
	$
ลายด้านซ้าย	L
ลายด้านขวา	R
จุดเริ่มต้น	.

รูปที่ 4.6 แสดงโค๊ตแทนลักษณะลาย

การที่พิจารณาว่าลายเริ่มต้น สิ้นสุดภายในหน้าต่างหรือไม่ โดยพิจารณาว่าเริ่มต้นในขอบเขตนี้หรือไม่

$$3 <= x_c <= 48$$

$$3 <= y_c <= 48$$

ถ้าเริ่มต้นภายในขอบเขตนี้ ถือว่า เริ่มต้นภายในหน้าต่าง นอกจากนั้น ถ้ามีการเริ่มต้นนอกขอบเขต ให้ถือว่า เป็นลายที่ต่อเนื่องกับหน้าต่างอื่น

4.2.6 การจำแนกลายนิ้วมือ

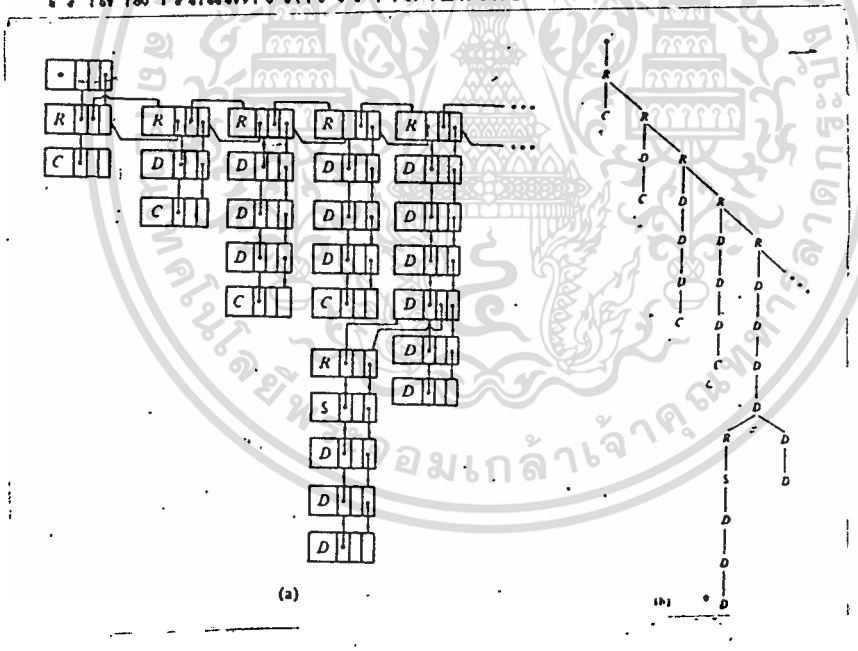
จากขั้นตอนที่แล้วเราเปลี่ยนข้อมูลในเหลือเพียงแค่ลักษณะลาย และลักษณะที่โดดเด่นของลาย ขั้นตอนนี้จะเอาข้อมูลจากขั้นตอนที่แล้วมาแทนลายในหน้าต่างนั้น โดยการแทนโครงร่างลายนิ้วมือทั้งหน้าต่าง โดยการแทนด้วยสัญญลักษณ์ที่สามารถแตกแขนง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ออกไปคล้ายต้นไม้ สามารถใช้อธิบายรูปภาพได้ และสามารถเขียนกลับมาเป็นโครงร่างแบบเดิมได้ การแทนโครงร่างนี้ จะประกอบด้วย 2 ส่วน คือ ส่วนของสัญญลักษณ์ที่แทนลักษณะโครงร่าง และเส้นเชื่อมความสัมพันธ์ระหว่างสัญญลักษณ์ที่สามารถบอกถึงลำดับความต่อเนื่องของโครงร่างได้

เส้นเชื่อมมีอยู่ 2 ลักษณะ คือ ในแนวดิ่ง หมายถึง โครงร่างที่เป็นโครงเดียวกัน และ ในแนวนอน หมายถึง โครงร่างใหม่ที่ไม่ใช่โครงร่างเดียวกัน แต่มีความสัมพันธ์กับโครงร่างเดิม ซึ่งสามรถบอกได้ว่าโครงร่างใหม่จะอยู่ด้านขวา หรือซ้ายของโครงร่างเดิม โดยดูได้จากสัญญลักษณ์ R หรือ L

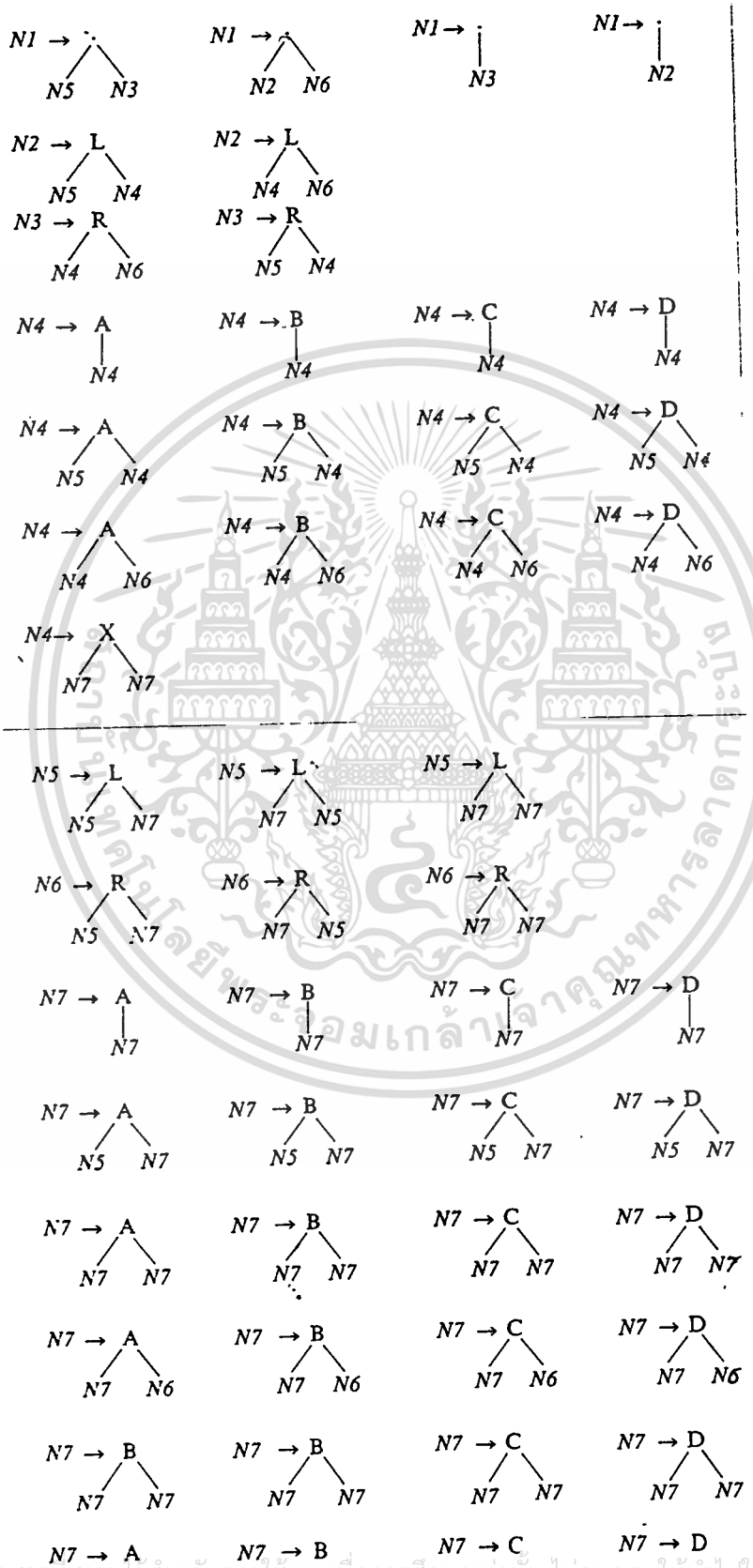
เราสามารถแสดงโครงร่างลายนิ้วมือ ได้ตามตัวอย่างดังนี้



รูปที่ 4.7 แสดงแผนภาพไบนารีทรี

จากที่กล่าวข้างต้นระบบจดจำลายนิ้วมือนี้จะพิจารณาที่ลักษณะโดดเด่น ข้อมูลอื่นๆที่ไม่เกี่ยวข้อง จึงเป็นสิ่งที่ไม่มีประโยชน์ ดังนั้นเราสามารถลดข้อมูลของลายนิ้วมือหนึ่งๆลงได้ด้วยการตัดข้อมูลที่ไม่มีประโยชน์ออก โดยใช้การไวยกรณ์แบบทรี ซึ่งมีลักษณะดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

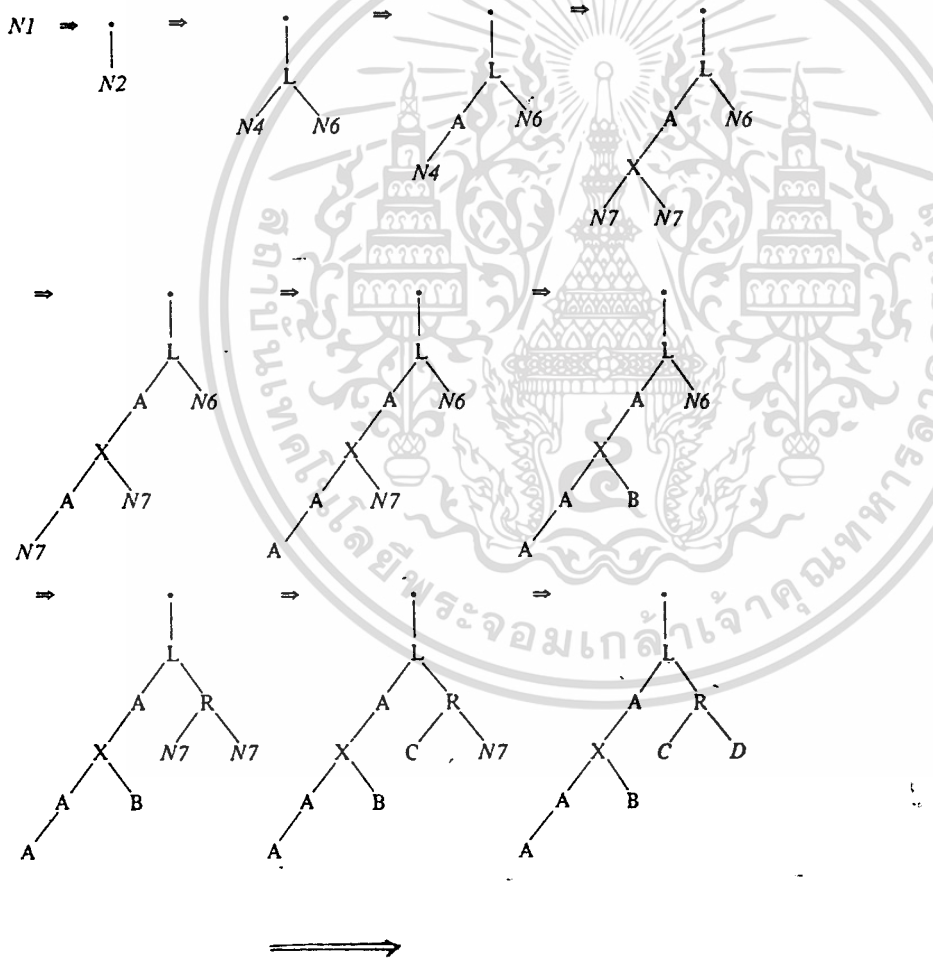


รูปที่ 4.8 รูปแบบไวยกรณ์แบบบริ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

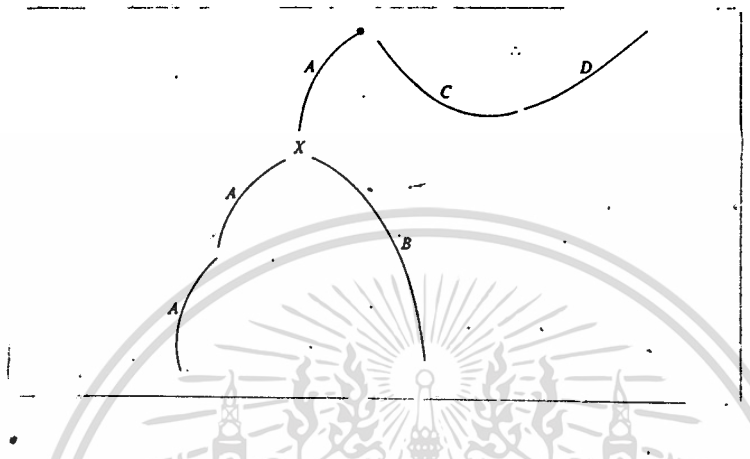
เราสามารถแทนข้อมูลที่ได้มาด้วยโค๊ต N1-N7 โดยจะแทนเมื่อพบว่าได้ข้อมูลตรงกับที่แสดงตามที่กำหนดในแต่ละส่วน ด้วยวิธีการนี้ก็จะสามารถลดข้อมูลลง ให้เหลือน้อยที่สุดเท่าที่จำเป็น

เราสามารถอธิบายการใช้ไวยกรณ์แบบทรี โดยใช้ตัวอย่างต่อไปนี้ ซึ่งเป็นการแปลงโค๊ตกลับมาเป็นโครงร่างของลายนิ้วมือ ดังรูปที่ 4.9 และจากโครงร่างลายนิ้วมือและสามารถเขียนลายนิ้วมือแบบที่แทนโดยโครงร่างนั้นได้ ดังรูปที่ 4.10



รูปที่ 4.9 ตัวอย่างการใช้ไวยกรณ์แบบทรีแทนโครงร่างลายนิ้วมือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.10 ลายนิ้วมือที่สามารถเขียนได้จากโครงร่างข้างต้น

ส่วนสุดท้ายของการจำแนกลายนิ้วมือคือการจดจำลายนิ้วมือ ทำได้โดยการเก็บข้อมูลลายนิ้วมือที่แทนโดยไวยกรณ์แบบทรี ข้อมูลของลายนิ้วมือที่แทนโดยใช้ไวยกรณ์แบบทรีทั้งหมด 16 หน้าต่างจะเป็นข้อมูลของแต่ละบุคคล การพิสูจน์ลายนิ้วมือจะทำโดยการนำข้อมูลลายนิ้วมือที่แทนโดยไวยกรณ์แบบทรีที่ต้องการเปรียบเทียบมาทำการเปรียบเทียบกันทีละหน้าต่างจนครบทั้ง 16 หน้าต่าง การที่จะยอมรับว่าลายนิ้วมือที่เปรียบเทียบนั้นเป็นลายนิ้วมือของบุคคลเดียวกันหรือไม่ ก็ต่อเมื่อ ข้อมูลลายนิ้วมือจะต้องเหมือนกันอย่างน้อย 14 ใน 16 หน้าต่าง

4.3 การทดลองและผลการทดลอง

จากทฤษฎีจากหัวข้อที่แล้ว สามารถนำมาพัฒนาเป็นโปรแกรมเพื่อใช้ในการทดลองนี้ โดยใช้ภาษาซีในการพัฒนา ตาสอัลกอริทึม ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. รับข้อมูลจากกล้อง
ขนาด 64 กิโลไบต์

2. เพิ่มความชัดเจน และ
ทำเป็นไบนารี่อิมเมจ

3. ผ่านขบวนการจัดการ
เบื้องต้น

4. ผ่านขบวนการกำจัด
ขั้นสุดท้าย

5. ผ่านขบวนการขยาย
คุณลักษณะ

6. ลดขนาดข้อมูล โดย
ใช้ไวยกรณ์แบบทรี

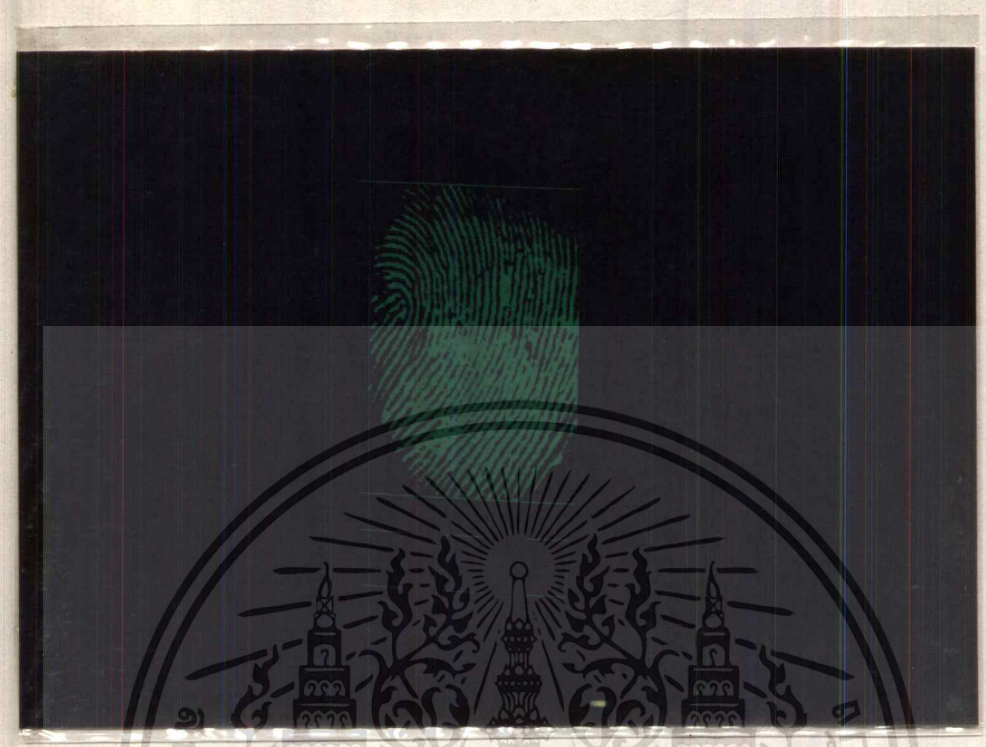
7. เก็บข้อมูลลงแผ่น
จานแม่เหล็ก

8. นำไปเปรียบเทียบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เราใช้ลายนิ้วมือจาก 5 บุคคลมาทำการทดลอง โดยมีบุคคลหนึ่งพิมพ์ลายนิ้วมือ 2 ครั้งเพื่อใช้ในการพิสูจน์หนึ่งลาย ที่เหลือใช้ในการจดจำลายนิ้วมือ ทำการวิเคราะห์และจดจำลายนิ้วมือทั้ง 5 แล้วเก็บลงแผ่นจานแม่เหล็ก จากนั้นทำการวิเคราะห์ลายที่เหลือแล้วนำมาเปรียบเทียบว่าตรงกับลายนิ้วมือที่มีอยู่ในฐานข้อมูล

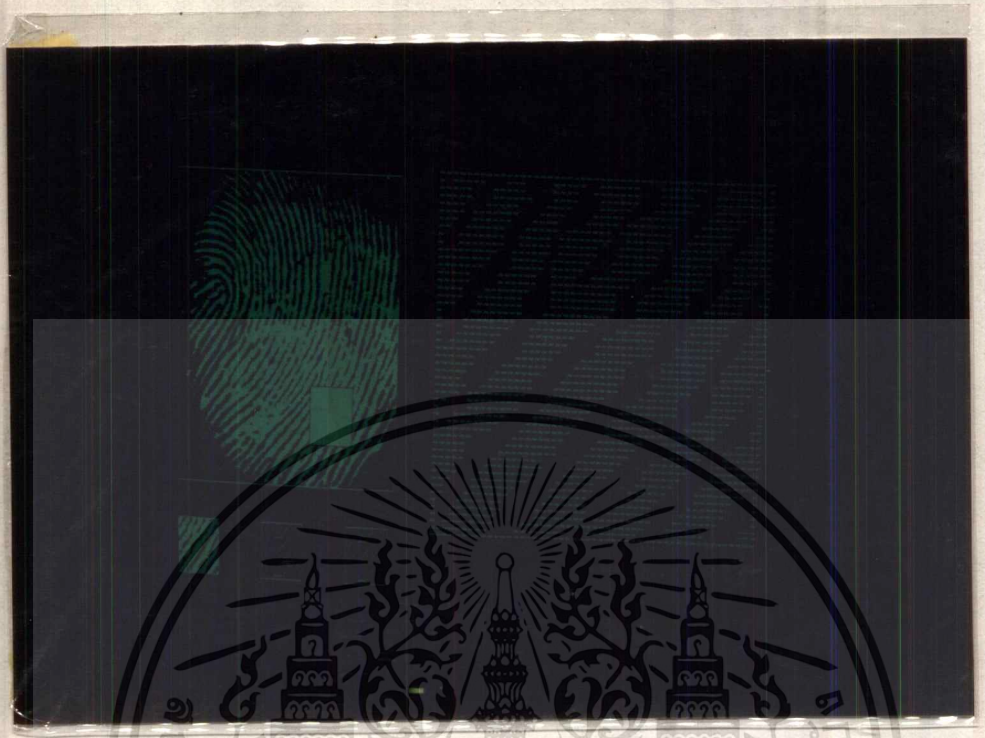
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.11 ภาพอินพุทที่รับมาจากกล้องวีดีโอ

รูปที่ 4.12 ตัวอย่างการแบ่งแยกหน้าต่าง (ด้านขวาคือภาพขยายของหน้าต่างนั้น)

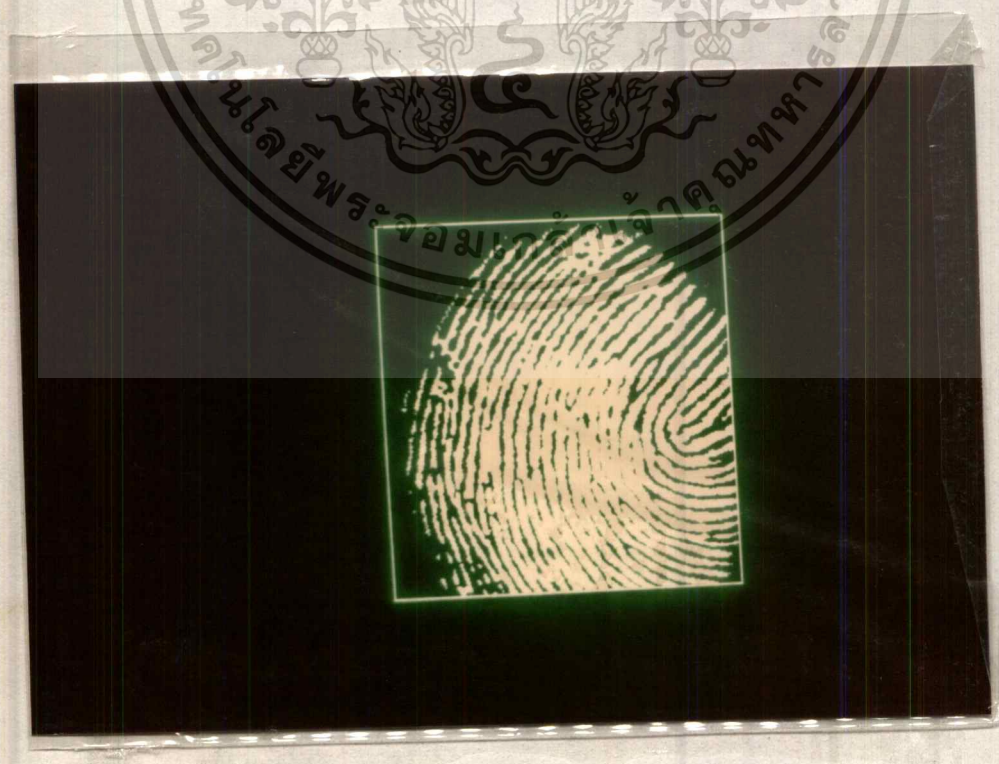
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.13 ภาพที่ผ่านขบวนการจัดการเบื้องต้น และขบวนการกำจัดขั้นสุดท้าย
(ภาพเล็กๆสองภาพตามลำดับ)

รูปที่ 4.14 ระหัสที่ได้หลังจากผ่านขบวนการกำจัดขั้นสุดท้าย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งไม่อนุญาตให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของผลงานทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 4.15 ตัวอย่างลายนิ้วมือบางตัวอย่างที่ใช้ในการทดลอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ผลการทดลอง

ในการทดลองเราทำการวิเคราะห์แล้วจำแนกลายนิ้วมือ 5 ลายเก็บเป็นฐานข้อมูล เก็บลงในแผ่นจานแม่เหล็ก แล้วนำลายนิ้วมืออีกลายของบุคคลที่มีลายนิ้วมืออยู่ในฐานข้อมูลนี้ มาทำการวิเคราะห์เพื่อการจดจำ

จากการทดลองเราได้ฐานข้อมูลของ 5 บุคคล แล้วทำการเปรียบเทียบ เพื่อทำการจดจำ โดยใช้ลายพิมพ์ลายนิ้วมือจากบุคคลที่มีลายนิ้วมืออยู่ในฐานข้อมูล จากนั้นนำลายนิ้วมือนั้นมาพิสูจน์ว่าเป็นลายเดียวกับลายที่มีอยู่ในฐานข้อมูลหรือไม่ ผลปรากฏว่าสามารถพิสูจน์ได้ว่าเป็นลายเดียวกับลายหนึ่งที่เก็บในฐานข้อมูล โดยใช้เวลาทั้งหมดในการค้นหา เพื่อทำการเปรียบเทียบกับฐานข้อมูลขนาด 5 บุคคล เป็นเวลา 45 วินาที

4.4 สรุปผลการทดลองและแนวทางการพัฒนา

จากการทดลองสามารถสรุปได้ว่า

1. เราสามารถลดขนาดของหน่วยที่ใช้เก็บลายนิ้วมือ จาก 256x256 เท่ากับ 64 กิโลไบต์ ให้เหลือเพียง 300-500 ไบต์เท่านั้น ขึ้นอยู่กับระหัสของแต่ละบุคคล

2. เวลาที่ใช้ในการแปลงระหัสข้อมูล ของแต่ละบุคคลประมาณ 40 วินาที (มีการแสดงระหัสของแต่ละหน้าต่างทางจอภาพด้วย ถ้าไม่แสดงเวลาที่ใช้จะลดลงอีก)

3. สำหรับฐานข้อมูลขนาด 5 บุคคล เราใช้เวลาในการค้นหาเพื่อเปรียบเทียบกับฐานข้อมูลนี้ประมาณ 45 วินาที

ข้อผิดพลาดในการทดลอง

1. เราไม่สามารถทำภาพอินพุทอย่างที่ต้องการในทางทฤษฎีได้ เพราะเราไม่มีอุปกรณ์ที่มีความสามารถเพียงพอ โดยในทฤษฎีที่ศึกษานั้นใช้สแกนเนอร์ (scanner) เป็นอุปกรณ์รับภาพอินพุท แต่ในการทดลองนี้เราจึงใช้กล้องวิดีโอในการรับภาพอินพุท

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยใช้การพิมพ์ลายนิ้วมือลงบนกระดาษ แล้วนำไปขยาย 16 เท่าจากปกติ แล้วจึงใช้กล้องถ่ายและบันทึกข้อมูลลงในแผ่นจานแม่เหล็ก ดังนั้น จึงทำให้ตำแหน่งของลายนิ้วมือที่ทำการบันทึกนั้นคลาดเคลื่อนไป ซึ่งเป็นข้อผิดพลาดที่อาจจะนำมาสู่การไม่สามารถจดจำลายนิ้วมือได้

2. เนื่องจากระบบจดจำลายนิ้วมือ เป็นระบบที่ใช้กับงานที่ละเอียดอ่อนมาก ระบบจึงต้องการความยืดหยุ่นค่อนข้างสูง ในการที่จะตรวจว่า เป็นลายนิ้วมือของบุคคลเดียวกันหรือไม่ ดังนั้นขบวนการที่ใช้ในงานจริง จึงต้องมีความละเอียดมากกว่าในการทดลองนี้ ซึ่งเป็นเพียงขั้นเริ่มต้นของการวิจัยและพัฒนาระบบจดจำลายนิ้วมือเท่านั้น

แนวทางในการพัฒนา

1. ควรจะหาวิธีการเพื่อลดขนาดของหน่วยความจำ ในการเก็บระหัสของแต่ละบุคคลให้ได้มากกว่านี้

2. ควรจะหาวิธีการเพื่อลดเวลา ในการแปลงระหัสของแต่ละบุคคลให้ได้มากกว่านี้

3. ควรจะหาวิธีการเพื่อลดเวลาในการค้นหาให้น้อยลง เนื่องจากมีความสำคัญมากในการใช้งานจริง เพราะในงานจริงจะมีข้อมูลจำนวนมาก ถ้าเราใช้เวลาในส่วนนี้มาก จะทำให้การค้นหาช้ายิ่งขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่5

ระบบเกี่ยวกับการมองเห็น

VISAUL SYSTEM

ในบทนี้ จะกล่าวถึงระบบการมองเห็นที่ใช้กล้อง 2 ตัว วางตำแหน่งต่างกัน ที่เรียกกันว่า ระบบสเตอริโอ โดยในหัวข้อ 5.1 กล่าวถึง รูปแบบของอิมเมจในระบบการมองเห็น หัวข้อ 5.2 เป็นการมองแบบสเตอริโอ หัวข้อ 5.3 เป็นคำแนะนำสำหรับการดูแบบสเตอริโอ หัวข้อ 5.4 เป็นการดูแบบสเตอริโอ ซึ่งจะกล่าวถึงการคำนวณในระบบสเตอริโอด้วย

5.1 รูปแบบของอิมเมจ (AN IMAGE MODEL) [4][5]

ในที่นี้คำว่า อิมเมจ จะหมายถึง ฟังก์ชัน ของความเข้มแสง 2 มิติ โดยจะมีรูปแบบดังนี้ f(x,y) ซึ่งขนาด หรือ แอมปิจูด (amplitute) ของฟังก์ชัน ที่พิกัดสปาเทียล (spatial coordinates) ที่จุด (x,y) จะแสดงค่าของความเข้มแสง (brightness) ของอิมเมจที่จุดนั้นๆ เมื่อแสงคือรูปแบบของพลังงานแบบหนึ่ง ดังนั้น f(x,y) จึงต้องมีขนาดไม่เท่ากับศูนย์ และจะต้องมีขนาดไม่เท่ากับอนันต์(finite) ดังนั้น จะได้ว่า

$$0 < f(x,y) < \infty \qquad \text{สมการที่ 1}$$

พื้นฐานโดยทางธรรมชาติของ f(x,y) อาจจะได้รับการพิจารณาเกี่ยวกับคุณสมบัติ 2 ประการ ดังต่อไปนี้คือ **ประการแรก** จะพิจารณาเกี่ยวกับแหล่งกำเนิดแสงทั้งหมดที่เกิดขึ้นเองบนฉาก ขณะที่**ประการที่สอง** จะพิจารณาเกี่ยวกับแสงสะท้อนทั้งหมด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ที่เกิดขึ้นจากการสะท้อนของวัตถุในฉากเอง คุณสมบัติทั้งสองประการนี้ จะถูกเรียกว่า การส่องสว่าง (illumination) และ การสะท้อน (reflectance) ซึ่งมีฟังก์ชันดังนี้คือ i(x,y) และ r(x,y) ตามลำดับ ทั้งสองฟังก์ชันนี้รวมกันจะกลายเป็นรูปแบบของ f(x,y) :

จะได้ว่า

$$f(x,y) = i(x,y).r(x,y) \qquad \text{สมการที่ 2}$$

เมื่อ

$$0 < i(x,y) < \alpha \qquad \text{สมการที่ 3}$$

และ

$$0 < r(x,y) < \alpha \qquad \text{สมการที่ 4}$$

จากสมการที่ (4) ชี้ให้เห็นถึงข้อเท็จจริงว่า ขอบล่างที่มีค่าเป็นศูนย์ "0" คือการดูดกลืนแสงทั้งหมด (total absorption) และขอบบนที่มีค่าเป็นหนึ่ง "1" คือการสะท้อนของแสงทั้งหมด (total reflectance) ที่เกิดขึ้น ดังนั้นธรรมชาติของฟังก์ชัน i(x,y) จะถูกกำหนดโดยแหล่งกำเนิดของแสง ขณะที่ฟังก์ชัน r(x,y) จะถูกกำหนดโดยคุณสมบัติของตัววัตถุในฉากนั้น ที่ความเข้มของฟังก์ชันอิมเมจแบบโมโนโครม (monochrome image) ที่จุด(x,y) จะถูกเรียกว่า ระดับสีเทา (gray level) หรือ (l) ของอิมเมจ ที่จุดนั้นๆจากสมการที่ (2) ถึง (4) จะเป็นหลักฐานว่า (l) ควรจะอยู่ในช่วงใด

จะได้ว่า

$$L_{min} <= 1 <= L_{max} \qquad \text{สมการที่ 5}$$

ตามทฤษฎีจะได้ว่า L_{min} จะมีค่าเป็นบวก และ L_{max} จะมีค่าเป็นจำกัด (finite) แต่ในทางปฏิบัติแล้ว $L_{min} = i_{min}.r_{min}$ และ $L_{max} = i_{max}.r_{max}$ และค่าของ L_{min} และ L_{max} ที่คาดหวังไว้ในการประยุกต์ใช้ในขบวนการทางอิมเมจในที่ร่มจะได้ค่าของ L_{min} มีค่าประมาณเท่ากับ 0.005 และ ค่าของ L_{max} มีค่า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประมาณเท่ากับ 100 ค่าต่างๆที่อยู่ในช่วงของ $[L_{min}, L_{max}]$ ถูกเรียกว่า ช่วงสีเทา (gray scale) ในทางปฏิบัตินี้ช่วงสีเทา จะเลื่อนมาอยู่ในช่วง [0,L] เพราะฉะนั้น จะได้ว่าเมื่อ I=0 จะได้รับการพิจารณาเป็น " มืด "(black) และเมื่อ I = L จะได้รับการพิจารณาเป็น " สว่าง " (white) ในฉากนั้นๆ

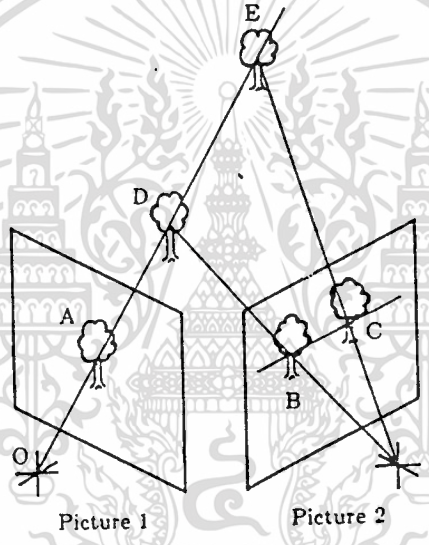
5.2 การมองแบบสเตริโอ (STEREO VISION) [4]

ปัญหาที่สำคัญของการทำความเข้าใจ อิมเมจทั่วๆไป คือต้องค้นพบในรูปแบบของ 3 มิติของวัตถุที่ได้มาจาก อิมเมจของภาพ 2 มิติ วิธีหนึ่งที่ใช้กัน คือต้องวัดขนาดของระยะทางจากกล้องถึงแต่ละจุดที่สนใจในฉาก ถ้าภาพของวัตถุที่เหมือนกันจากตำแหน่ง 2 ตำแหน่งที่มีความแตกต่างกัน พวกเราก็สามารถที่จะวัดระยะทาง ซึ่งห่างจากกล้องถึงวัตถุโดยใช้วิธีการของสามเหลี่ยม (triangulation) วิธีการนี้ถูกรู้จักในฐานะที่เป็นสเตริโอ (stereo) หรือ การมองแบบไบโนคูลา (binocular vision) วิธีการนี้เป็นแหล่งพื้นฐานของข้อมูลสำหรับมนุษย์ เกี่ยวกับวัตถุ 3 มิติ ภาพในสภาพแวดล้อมนั้นๆ

ภาพของไบโนคูล ของกล้องทั้งคู่ได้ถูกแสดงไว้ในรูปที่ 1 อิมเมจของวัตถุในภาพที่ได้จากกล้องทั้งคู่จะถูกทำเป็นรูปแบบ โดยเป็นรูปแบบรังสีของแสงโดยเริ่มต้นที่วัตถุแล้วผ่านทะลุไปยังจุดตรงกลางของเลนส์ (lens) ในทางกลับกันจุดตรงกลางของเลนส์และจุดในอิมเมจ จะกำหนดเส้นตามยาวซึ่งวัตถุต้องตกลงที่ฉาก เส้นตามยาวของวัตถุนี้จะถูกวางไว้ในตำแหน่งไหน และมันจะอยู่ห่างออกไปเท่าไหร่ สิ่งนี้จะถูกกำหนดอย่างง่ายดาย กับอิมเมจที่ 2 ซึ่งได้จากกล้องอีกตัวหนึ่ง ถ้าเราสามารถหาจุดในอิมเมจ ที่ 2 ซึ่งมีคุณสมบัติเหมือนกัน หรือที่เกี่ยวโยงกันถึงจุดใน อิมเมจตัวที่ 1 ในตอนนั้นจุดนั้นและศูนย์กลางของเลนส์ของมัน จะกำหนดรังสีของแสงที่ 2. ได้ ถ้าจุด 2 จุดเกิดจากวัตถุที่เหมือนกัน และถ้ารู้ตำแหน่งของกล้องได้ถูกต้อง และถ้ากล้องทั้ง 2 เป็นเชิงเส้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(linear) ตอนนั้นวัตถุต้องถูกวางไว้ที่จุด ซึ่งจุดนั้นจะต้องมีรังสี 2 รังสีซึ่งจะคำนวณเพื่อที่จะหาจุดตัดกัน



รูปที่ 5.1 การมองเห็นของภาพ

ปัญหาของข้อมูลความลึก จากคู่สเตริโอของอิมเมจ จะแบ่งออกได้เป็น 4 ส่วน คือการหาลักษณะสำคัญ (features) ของอิมเมจ ซึ่งถูกจดจำได้อย่างง่ายในอิมเมจ ทั้งคู่ การเข้าคู่กัน (matching) ของลักษณะสำคัญนี้ใน 2 อิมเมจ การกำหนดความสัมพันธ์ของตำแหน่งกล้อง และข้อวินิจฉัยเกี่ยวกับระยะทางที่ห่างจากกล้องจน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ถึงวัตถุซึ่งจะคำนวณได้จากลักษณะสำคัญในอิมเมจทั้งสอง

การรู้ข้อมูลที่แน่นอนของรูปแบบของกล้อง ซึ่งจะทำให้สามารถค้นหา การเข้าคู่กันสำหรับลักษณะสำคัญที่ได้รับมาในอิมเมจหนึ่งๆ ซึ่งจะถูกจำกัดบนเส้นอื่นๆมากกว่าบนอิมเมจทั้งหมด จากรูปที่ 5.1 แสดงว่าวัตถุจะรับผิดชอบสำหรับลักษณะสำคัญที่จุด A. ในอิมเมจทางด้านซ้าย ซึ่งต้องวางเป็นทางยาวไปตามจุดเริ่มต้นของรังสี ที่จุดกำเนิดของระบบพิกัด (coordinate) ทางงด้านซ้ายที่จุด "O" และการปฏิบัติโดยผ่านทะลุลักษณะสำคัญที่จุด A. อิมเมจของวัตถุในภาพทางด้านขวา ต้องวางนอนลงตามความยาวของการฉายของรังสี OA ลงบนระนาบ (plane) ของอิมเมจทางด้านขวา ซึ่งก็คือเส้นตรง BC ซึ่งเส้นตรงนี้จะถูกเรียกว่า "เส้นอีพิโพลา" (epipolar line)

มีข้อสังเกตว่า ทุกๆจุดบนเส้นอีพิโพลาที่ได้รับมาใน อิมเมจทางด้านขวา ต้องจับเข้าคู่กับ บางจุดบนตามความยาวของเส้นอีพิโพลาที่ตรงกัน(corresponding) กับในอิมเมจทางด้านซ้าย และในทำนองเดียวกัน แต่เป็นในทางกลับกัน ทั้งๆที่เราไม่ทราบพารามิเตอร์ของกล้องที่แน่นอน การค้นหาการเข้าคู่กันยังคงถูกจำกัดในพื้นที่แคบๆตามแนวความยาวบนเส้นอิพิโพลา (epipolar)

5.2.1 การหาคุณลักษณะสำคัญ (finding features)

วัตถุประสงค์ของการหาคุณลักษณะสำคัญ ในอิมเมจหนึ่งๆ ซึ่งต้องไม่มีความไม่เด่นชัดเกิดขึ้น ในการเข้าคู่กันในอิมเมจอื่นๆ ตัวอย่างในแต่ละพิเซล (pixel) จะมีการบรรจุข้อมูลน้อย ในความเป็นจริงคือว่า พิเซลสองพิเซลในสองอิมเมจจะมีระดับสีเทา ที่เหมือนกันจะเป็นหลักเล็กน้อย ซึ่งพิเซลเหล่านี้จะถูกผลิตขึ้นโดย วัตถุที่เหมือนกันในฉาก การเข้าคู่กันที่ไม่มีความไม่เด่นชัดเกิดขึ้น จะถูกพบมากในบริเวณของรูปซึ่งมีการบรรจุข้อมูลมาก หรือ ความแปรปรวนสูง (high varience) หน่วยปฏิบัติการของโมราแวค (moravec's interest operator) คือเทคนิคสำหรับการหาบริเวณ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เช่นนี้ โดยครั้งแรกจะคำนวนผลบวกกำลังสองของ ความแตกต่างระหว่างแต่ละพิเซล ในหน้าต่าง (window) หนึ่งๆ และพิเซลรอบข้าง (pixel's neighbors) ในแต่ละทิศทางซึ่งมี 4 ทิศทางดังนี้คือ แกนนอน แกนตั้ง และทแยงสองทิศทาง (two diagonals) ผลลัพธ์ในการบวกทั้ง 4 ทิศทาง และเกณฑ์ที่น่าสนใจสำหรับแต่ละพิเซล คือค่าต่ำสุดของผลบวก จุดที่น่าสนใจจะอยู่ที่นี่ เกณฑ์นี้จะได้รับค่าสูงสุดเฉพาะแห่ง ดังนั้น จุดที่น่าสนใจจะมีความแปรปรวนสูงในทิศทางทั้งหมด และต้องค้นหาการเข้าคู่กัน โดยมีความเด่นชัดเกิดขึ้นในอิมเมจอื่นๆ

แหล่งที่ดีอีกแหล่งหนึ่งของจุดต่างๆ สำหรับการเข้าคู่กัน ซึ่งมีความเด่นชัดเกิดขึ้น จะอยู่ที่ขอบของวัตถุ เพราะว่าขอบของวัตถุที่จุดต่างๆนี้ จะมีลักษณะไวต่อการเปลี่ยนแปลงความเข้มแสง สำหรับสเตริโอ และขอบของวัตถุอาจจะถูกตัดทอน โดยวิธีการตรวจขอบของวัตถุทั่วๆไป

รูปแบบการคำนวนของ ระบบการมองแบบสเตริโอของมนุษย์ (human stereo vision) จะใช้จุดตัดแกนศูนย์ (zero-crossing) ในอิมเมจลาปลาสเซียล หลังจากผ่านฟิลเตอร์ แบบกรองความถี่ต่ำของเกาส์เซียน ซึ่งจะทำให้ข้อมูลเรียบขึ้น (smoothing) จุดตัดแกนศูนย์ คือจุดที่ผ่านการทำอนุพันธ์อันดับสองของเครื่องหมายการเปลี่ยนแปลงอิมเมจ

5.2.2 การเข้าคู่กัน (matching)

การเข้าคู่กัน คือขบวนการค้นหา ซึ่งแบ่งออกเป็น 2 ส่วน คือ การวัดความแตกต่าง (difference measure) ,แทนการค้นหา (search strategy)

การวัดความแตกต่าง มีอยู่ 2 ชนิด ของการวัดความแตกต่าง จะถูกใช้เพื่อประเมินการเข้าคู่กันระหว่างลักษณะสำคัญที่ตรงกันของ 2 อิมเมจ ชนิดแรก เหมาะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับการใช้เมื่อลักษณะสำคัญ เป็นพื้นที่ที่ประกอบด้วยบริเวณต่างๆ อีกชนิดหนึ่งจะถูกใช้สำหรับลักษณะสำคัญที่เป็นเส้น เช่น ขอบ (edges)

การวัดแบบธรรมดาของความแตกต่างระหว่างลักษณะสำคัญของ 2 พื้นที่ คือ ผลบวกของกำลังสองของผลต่างระหว่างพิเซลที่ตรงกัน เรียกว่า L_2

จะได้ $$L_2 = \Sigma(E_1 - E_2)^2$$ สมการที่ 5

เมื่อ E_1 และ E_2 คือค่าพิเซลของ 2 อิมเมจ และที่รวมกันจะขยายออกไปตามขอบเขตของความสนใจ การวัดนี้ต้องการคูณกันสำหรับแต่ละพิเซลในขอบเขตการวัดของผลรวมของค่าสมบูรณ์ของผลต่างของ 2 พื้นที่ เรียกว่า L_1

จะได้ $$L_1 = \Sigma|(E_1 - E_2)|$$ สมการที่ 6

สมการนี้ต้องการเพียงแต่ค่าสมบูรณ์เท่านั้น

การวัดทั้งสองนี้ต้องการความไวในการเปลี่ยนแปลงในความแตกต่างและความสว่างระหว่างอิมเมจ ความแตกต่างนี้เกิดขึ้นเมื่อทั้งสองภาพของวัตถุที่เหมือนกันจะผ่านขบวนการที่แตกต่างกัน เมื่อได้รับอิมเมจที่ถูกแยกมาใต้สภาพของแสงที่แตกต่างกัน และเนื่องจากความสว่างที่ปรากฏขึ้นบนผิวหน้าจะเปลี่ยนแปลงกับมุมเดิม การวัดที่เหมือนกันระหว่าง 2 อิมเมจ ซึ่งไม่มีความไวต่อการเปลี่ยนแปลงความแตกต่างและความสว่าง ซึ่งเป็นความเกี่ยวพันกันแบบปกติ (normalized correlation) คือ

$$\Sigma (E_1 . E_2) / \sqrt{} (\Sigma E_1^2 . \Sigma E_2^2)$$

คุณลักษณะที่เป็นเส้น (linear) ที่คล้ายคลึงกัน ถ้าคุณลักษณะพวกนี้มีทิศทางเดียวกัน และถ้าเกรเดียนต์ความเข้ม (intensity gradient) ของมัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เหมือนกัน เพื่อที่จะพิจารณาขอบ 2 ขอบที่เข้าคู่กัน ซึ่งการเข้าคู่กันจะสนใจเพียงเกรเดียนต์ความเข้มของขอบที่มีเครื่องหมายเหมือนกัน

แผนการค้นหา การหาจุดร่วมใน 2 อิมเมจเป็นปัญหาที่ยากมากของการแปลความหมายอิมเมจสเตริโอ ระบบการหาคุณลักษณะสำคัญที่ดีที่สุด จะยังคงเลือกคุณลักษณะสำคัญ ซึ่งการเข้าคู่กันจะปรากฏคุณลักษณะสำคัญมากกว่าหนึ่งคุณลักษณะ ในอิมเมจอื่นๆ จากข้อเท็จจริงที่ว่า ถ้าฉากมีรูปแบบที่ซ้ำกัน ซึ่งเหมือนกับหน้าต่างของตึก จะทำให้การจับคู่กันจะมีความไม่เด่นชัดเกิดขึ้น อย่างหลีกเลี่ยงไม่พ้น ยิ่งไปกว่านั้น เนื่องจากการปิดบังบริเวณนี้จะมีคุณลักษณะสำคัญ ในอิมเมจหนึ่งๆ ซึ่งจะไม่ปรากฏในอิมเมจอื่นๆ มันเป็นไปไม่ได้ที่จะหาการเข้าคู่กันสำหรับคุณลักษณะสำคัญนี้

การจับคู่กันในบริเวณพื้นที่ขนาดใหญ่ จะทำให้เกิดความซับซ้อน ในความคิดที่เราหาคุณลักษณะสำคัญที่เป็นพื้นที่ N ในอิมเมจหนึ่งๆ ในแต่ละพื้นที่จะมีพื้นที่เฉลี่ยของ w^2 พิเซล ในกรณีที่เลวร้ายที่สุด คือพวกเราไม่ทราบได้เลยว่าตำแหน่งต่างๆของกล้องซึ่งมันมีประโยชน์ ในการหาตำแหน่งต่างๆของพิเซลในอิมเมจอื่นๆ ซึ่งจุดศูนย์กลางของพิเซลในคุณลักษณะสำคัญของพื้นที่ ซึ่งมีลักษณะเดียวกัน มันจะมีความเกี่ยวพันกันกับการหา w^2 พิเซล จากตำแหน่งพิเซลแต่ละตำแหน่งในอิมเมจอื่นๆ สำหรับแต่ละคุณลักษณะสำคัญ N เราต้องการ NW^2L^2 เพื่อเปรียบเทียบ เมื่อ L^2 คือ พื้นที่ของอิมเมจอื่นๆ

มีคุณสมบัติ 2 ประการของวัตถุคือ **การเกาะรวมกัน** (cohesiveness) และ**ความทึบแสง** (opacity) ซึ่งสามารถช่วยในการตัดสินใจระหว่างการเข้าคู่กันอย่างไม่เด่นชัดเกิดขึ้น ที่เป็นเช่นนี้ เนื่องจากการเกาะรวมกันของวัตถุ คือ ระยะจากกล้องถึงวัตถุ ซึ่งปรากฏอยู่ใกล้กันและกัน ในอิมเมจมีความเหมือนกันโดยประมาณ เนื่องจากความทึบแสงของวัตถุแต่ละจุดในแต่ละอิมเมจ จะมีความลึกเพียงหนึ่งเดียวที่สัมพันธ์กับมัน ดังนั้นแต่ละจุดในแต่ละอิมเมจ จะมีการเข้าคู่กันอย่างมากที่สุดในอิมเมจอื่นๆ

ถ้าเราสามารถทราบพารามิเตอร์ของกล้อง ดังนั้นการเข้าคู่กันก็สามารถทำได้ง่ายยิ่งขึ้นโดยการนำเอาประโยชน์จากข้อเท็จจริงที่ว่า การเข้าคู่กันจะเกิดขึ้นเพียง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ที่ตามแนวความยาวบนเส้นอิพิโพลา สิ่งนี้เป็นสิ่งสำคัญที่พิเศษสุดในรูปแบบที่จะคำนวนของระบบการมองแบบสเตริโอ ซึ่งการค้นหาจะถูกจำกัดขอบเขต ตามแนวความยาวบนเส้นแกนนอนของตำแหน่งทางแกนตั้งที่เหมือนกันในอิมเมจทั้งคู่

สำหรับปัญหาของสเตริโอที่เป็นพื้นฐานของขอบ คือ ตามแนวความยาวบนเส้นอิพิโพลาของคุณลักษณะสำคัญที่จะเข้าคู่กันในอิมเมจ กับเส้นอิพิโพลาที่ตรงกันตามความยาวของคุณลักษณะสำคัญในอิมเมจอื่นๆ คุณลักษณะสำคัญบางอย่างอาจจะไม่มีการจับเข้าคู่กัน เนื่องจากเกิดการปิดบังกันเกิดขึ้น ข้อบังคับที่ใช้มีอยู่ว่าคุณลักษณะสำคัญที่ใกล้เคียงกัน ในอิมเมจหนึ่งๆ จะทำการจัยคู่กับคุณลักษณะสำคัญในอิมเมจอื่นๆ และข้อกำหนดของคุณลักษณะสำคัญที่เข้าคู่กันจะถูกสงวนตามแนวความยาวบนเส้นอิพิโพลา

5.2.3 การแก้ไขสำหรับรูปแบบของกล้อง
(solving for the camera model)

ข้อมูลเกี่ยวกับตำแหน่งที่มีความสัมพันธ์ของกล้องทั้งสองตัวนี้ ถูกเรียกว่ารูปแบบของกล้อง (camera model) ถ้าพารามิเตอร์ของกล้องถูกรู้อย่างแน่นอนตอนนั้นจุด ซึ่งมีลักษณะตรงกันถึงจุดที่ได้รับในอิมเมจทางด้านซ้ายไปตามแนวความยาวบนเส้นอิพิโพลาในอิมเมจทางด้านขวา ตำแหน่งของเส้นอิพิโพลาขึ้นอยู่กับพารามิเตอร์ของกล้อง และถ้าพารามิเตอร์เหล่านี้ไม่แน่นอนจุดที่เข้าคู่กัน จะวางลงในตำแหน่งต่างๆบนเส้นอิพิโพลา ดังนั้นระยะทางระหว่างจุดเข้าคู่กันกับเส้นอิพิโพลา คือการวัดความผิดพลาดของพารามิเตอร์ของกล้องในขณะใดขณะหนึ่ง

5.3 คำแนะนำสำหรับการดูแบบสเตริโอ (INTRODUCTION STEREOSCOPY)

สเตริโอสโคปี้ (STEREOSCOPY) เป็นเทคนิคขั้นพื้นฐานสำหรับการคำนวน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ระยะของข้อมูล เกี่ยวกับฉาก ระบบภาพสเตริโอที่ เป็นนอนคอนเวอร์เจนท์ (nonconvergent) ประกอบด้วยกล้องที่ไม่เคลื่นที่ทั้งสองตัว ซึ่งกล้อง 2 ตัว ถูกแยกออกจากกันโดยระยะทางของเส้นเบส (baseline) และให้แกนของเลนส์ขนานกันทั้งสองกล้อง การทำการติดตั้งกล้องทั้งสองตัวเพื่อที่จะอิมเมจซ้ายและอิมเมจขวามานั้นจะสำเร็จ โดยการคำนวนการเคลื่อนที่หรือความแตกต่างระหว่างจุดที่เป็นลักษณะสำคัญที่ตรงกัน 2 จุด ในอิมเมจทางซ้ายและอิมเมจขวา ซึ่งจะทำให้คู่ลำดับทาง 3 มิติของจุดอิมเมจในฉากสามารถจะถูกค้นพบได้

ในการออกแบบในระบบสเตริโอ จะเกี่ยวข้องกับการเลือก พารามิเตอร์ต่างๆ ซึ่งพารามิเตอร์ต่างๆนี้จะประกอบด้วย ความยาวโฟกัสของกล้องทั้งสอง ช่วงระยะเวลาการแชมปิ้งของอิมเมจ ระยะห่างของเส้นเบส และ ระยะทางจากกล้องทั้งสองถึงฉาก เพื่อที่จะทำให้สามารถทราบการจับเข้าคู่กันของจุดต่างๆ ที่เป็นคุณลักษณะสำคัญได้แน่นอน และเพื่อที่จะหลีกเลี่ยง การถูกบังเท่าที่จะหลีกเลี่ยงได้ ดังนั้นผลคูณของความยาวของเส้นเบส กับความยาวโฟกัสจะต้องมีค่าน้อย อย่างไรก็ตามการประมาณระยะที่แม่นยำจะต้องทำให้ผลคูณของทั้งสองนี้มีค่ามาก การประมาณระยะที่แม่นยำสามารถปรับปรุงได้โดยการลดช่วงเวลาการแชมปิ้งลง แต่สิ่งนี้โดยทั่วไปจะถูกบังคับโดยข้อจำกัดทางฟิสิกส์ของอุปกรณ์ของอิมเมจ ดังนั้นการตัดสินใจว่าจะให้การประมาณระยะมีค่าความผิดพลาดเท่าไหร่ถึงจะยอมรับได้ และตอนนั้นก็จะเตรียมที่จะเลือกการออกแบบพารามิเตอร์เพื่อที่จะให้ได้พบกับสิ่งที่ต้องการได้

5.4 การดูแบบสเตริโอ (STEREOSCOPY)

จากรูปที่ 2 ซึ่งแสดงให้เห็นถึง รูปแบบพินโอล์ (pinhole) ของระบบสเตริโอนอนคอนเวอร์เจนต์ ซึ่งแต่ละระนาบของอิมเมจ จะได้รับผลกระจายเกี่ยวกับจุดโฟกัสของตัวมันเอง รูปแบบของพินโอล์ จะแตกต่างจากรูปแบบของเลนส์ในแต่ละจุดของ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(similar triangles) ดังนั้นจะได้สมการของ การฉายแบบเพอร์สเปกทีฟ (perspective) จะถูกค้นพบดังนี้

$$X_L = f.X / Z \qquad \text{สมการที่ 7}$$

$$X_R = f(X-B) / Z \qquad \text{สมการที่ 8}$$

$$Y_L = Y_R = f.Y / Z \qquad \text{สมการที่ 9}$$

ทั้ง 3 สมการข้างบนนี้สามารถถูกแปลงกลับไปเพื่อแก้ให้อยู่ในรูปของ X,Y,Z โดยกำหนดความแตกต่างไว้ดังนี้ คือกำหนดให้

$$d = X_L - X_R$$

ดังนั้นสมการของการฉายแบบเพอร์สเปกทีฟ จะถูกเขียนให้ใหม่ได้ดังนี้

$$X = b.X_L / d \qquad \text{สมการที่ 10}$$

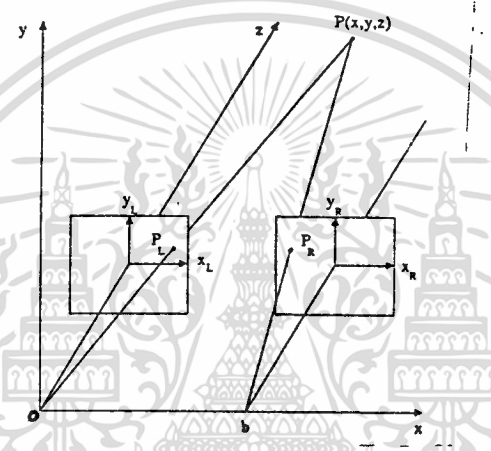
$$Y = b.Y_L / d \qquad \text{สมการที่ 11}$$

$$Z = b.f / d \qquad \text{สมการที่ 12}$$

สมการต่างๆนี้จะจัดการขั้นพื้นฐานสำหรับการได้มาซึ่งโครงสร้างที่เป็น 3 มิติ จากอิมเมจสเตริโอ ส่วนใหญ่อัลกอริทึมสเตริโอ (stereo algorithms)จะประกอบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัววัตถุนั้น โดยไม่คำนึงถึงระยะทางของมันจากผู้สังเกต แต่ละจุดทำการฉายถึงจุดเดียวบนระนาบของอิมเมจ สำหรับวัตถุประสงค์ของการวิเคราะห์สิ่งนี้ จะรู้ว่ารูปแบบพินโฮล์จะจัดหาการประมาณที่เป็นเหตุเป็นผลของรูปแบบของเลนส์



รูปที่ 5.2 แสดงการดูแบบสเตอริโอ

จากการแสดงในรูปที่ 2 จะได้ว่าพิกัดของจริง (world coordinate) ของ 3 มิติ คือ (X,Y,Z) ระบบพิกัดของอิมเมจทางด้านซ้าย คือ (X_L,Y_L) มีจุดกำเนิดของมันที่พิกัดของจริง คือ (0,0,f)และจุดของโฟกัสทางด้านซ้ายที่ตรงอยู่ที่ (0,0,0) ส่วนระบบพิกัดของอิมเมจทางด้านขวา คือ (X_R,Y_R) มีจุดกำเนิดของมันอยู่ที่ (b,0,f) และจุดโฟกัสที่อยู่ทางด้านขวาที่ตรงกันอยู่ที่ (b,0,0) ในการคำนวณนี้จะกำหนดให้ b คือ ระยะเส้นเบส และ f คือ ความยาวของโฟกัส

แต่ละจุดในส่วนของภาพจะถูกฉายลงมาเป็นรังสี ผ่านทะลุจุดโฟกัสไปลงบนระนาบของอิมเมจของทั้งทางด้านซ้าย และทางด้านขวา ให้จุดพิกัดของ 3 มิติทั่วๆไป คือ P(X,Y,Z) ซึ่งถูกแสดงในรูปที่ 2 โดยการพิจารณาหลักการของสามเหลี่ยมคล้าย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ด้วยกัน 3 ส่วนคือ

1.การขยายคุณลักษณะสำคัญ (feature extraction)

2.การเข้าคู่กันที่เป็นลักษณะสำคัญ (feature matching)

3.การมีโครงสร้างใหม่ของ 3 มิติ (3 dimensional reconstruction)

สำหรับแต่ละคู่ ของจุดของคุณลักษณะสำคัญ ซึ่งได้รับการเข้าคู่กัน ความไม่เหมือนกัน จะถูกคำนวณ และพิกัดแบบ 3 มิติ จะถูกสร้างขึ้นมาใหม่ โดยการใช้สมการ (4) - (6)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก

กิตติกรรมประกาศ

โครงงาน และ ปริญญานิพนธ์นี้มีอาจที่จะสำเร็จลงได้ตามวัตถุประสงค์ถ้าขาดความร่วมมือและคำแนะนำต่างๆ ดังนั้นในนามของผู้จัดทำจึงต้องขอขอบคุณแก่ผู้ให้ความร่วมมือทุกฝ่าย อาทิ อาจารย์ เพื่อนๆที่ได้ช่วยในการให้ข้อคิดเห็น ตลอดผู้ร่วมงานในกลุ่มทุกคนที่ได้ช่วยกันแก้ปัญหา และสุดท้ายคือ อาจารย์ เกษตร์ ศิริสันติสัมฤทธิ์ ที่ได้ให้ความช่วยเหลือทุกด้านในการทำวิจัยครั้งนี้อย่างมาก ตลอดจนจัดสรรหาเครื่องมือในการทดลอง และหวังว่าความดีที่ได้จากปริญญานิพนธ์นี้ ขอมอบกลับสู่ผู้ที่ให้ความร่วมมือ และกำลังใจทุกท่าน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หนังสืออ้างอิง

1. R. Gonzalez and P. Wintz,"Digital Image Processing," Addison-Wesley Publishing Company,Inc.,1987.

2. A. Huertas and G. Medioni, " Detection of Intensity Changes with Subpixel Accruracy Using Laplacian-Guassian Masks," Pattern Recognition,vol. 8,Sep. 1986.

3. B. Moayer and K.S. Fu, " A Tree System approach for Fingerprint Pattern Recognition,"Pattern Recognition vol. 6,May 1976.

4. A. Barr, P. Cohen and E. Feigenbaum, " The Handbook of Artificial Intellegence Volume 3," Addison-Wesley Publishing Company,Inc.,1982.

5. J. Rodriguez and J. Aggarwal,"Stochastic Analysis of Stereo Quantization Error ," Pattern Anal. Machine Intell.,vol. 9, Nov. 1990.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้