สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

ระบบจดจำตัวอักษรคำบรรยายภาพยนตร์

Movie Subtitle Character Recognition System

โดย

นางสาว ณิชจิรา ชอบธรรม

นาย ธนพล จึงเจริญสุขยิ่ง

นาย ธนพัฒน์ วรรณชาติ

วพ.
ว 4218
ฉบ.1

เลขหมู่.....
เลขทะเบียน.....72664.....
วัน,เดือน,ปี.....21 มิ.ย. 2550.....

.b.....11717045.....
.i.....

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

ภาควิชาวิศวกรรมสารสนเทศ

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2549

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Movie Subtitle Character Recognition System

By

Miss. Nichira Choubtum

Mr. Thanapol Juengjaroensukying

Mr. Thanapat Wanachart

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENT FOR THE DEGREE OF

BACHELOR IN DEPARTMENT OF INFORMATION ENGINEERING

FACULTY OF ENGINEERING

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

2006

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อปริญญานิพนธ์	ระบบจดจำตัวอักษรคำบรรยายภาพยนต์	
ชื่อนักศึกษา	นางสาว ณิชจิรา ชอบธรรม	รหัสประจำตัว 46010229
	นาย ธนพล จึงเจริญสุขยิ่ง	รหัสประจำตัว 46010276
	นาย ธนพัฒน์ วรรณชาติ	รหัสประจำตัว 46010278
อาจารย์ที่ปรึกษา	อ. ภุชงค์ หงษ์สุวรรณ	
ระดับการศึกษา	ปริญญาตรี วิศวกรรมศาสตรบัณฑิต	
	สาขาวิศวกรรมสารสนเทศ	
ภาควิชา	วิศวกรรมสารสนเทศ	
ปีการศึกษา	2549	

ปริญญานิพนธ์นี้ได้รับความเห็นชอบจากอาจารย์ที่ปรึกษาเป็นที่เรียบร้อยแล้ว

.....

อ. ภุชงค์ หงษ์สุวรรณ

อาจารย์ที่ปรึกษา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อปริญญานิพนธ์	ระบบจดจำตัวอักษรคำบรรยายภาพยนต์		
ชื่อนักศึกษา	นางสาว ณิชจิรา	ชอบธรรม	รหัสประจำตัว 46010229
	นาย ธนพล	จึงเจริญสุขยิ่ง	รหัสประจำตัว 46010276
	นาย ธนพัฒน์	วรรณชาติ	รหัสประจำตัว 46010278
อาจารย์ที่ปรึกษา	อ. ภุชงค์	หงษ์สุวรรณ	
ระดับการศึกษา	ปริญญาตรี วิศวกรรมศาสตรบัณฑิต		
	สาขาวิศวกรรมสารสนเทศ		
ภาควิชา	วิศวกรรมสารสนเทศ		
ปีการศึกษา	2549		

บทคัดย่อ

ทุกวันนี้ความต้องการทางด้านมัลติมีเดียเป็นอีกปัจจัยหนึ่งที่มีความสำคัญในการดำรงชีวิตอยู่ในสังคมปัจจุบัน โปรแกรมนี้ถูกสร้างขึ้นมาเพื่ออำนวยความสะดวกในการใช้งานทางด้านการรับชมภาพยนตร์ที่มีการใช้คำบรรยายภาพยนตร์ โดยจะอำนวยความสะดวกในการคัดลอก ดัดแปลงคำบรรยายภาพยนตร์ ที่มีอยู่แล้วในภาพยนตร์เพื่อนำมาใช้ในการทำงานต่อไป โดยจะสามารถดักจับคำบรรยายภาพยนตร์ที่ออกมาเปลี่ยนเป็นข้อมูลตัวอักษรด้วยกระบวนการทางด้านการประมวลผลรูปภาพเพื่อตรวจสอบข้อความสัญลักษณ์โดยมีความแม่นยำทั้งประโยคและระยะเวลาของคำบรรยายภาพยนตร์ จากนั้นจะทำการแปลงข้อมูลที่ได้เป็นไฟล์คำบรรยายภาพยนตร์แบบเอเอสเอส(.ass) ซึ่งสามารถนำไปใช้กับโปรแกรมซับไตเติ้ลสตูดิโอ(Subtitle Studio)ทั่วไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Thesis Title	Movie Subtitle Character Recognition System	
Student	Miss. Nichira Choubtum	ID. 46010229
	Mr. Thanapol Juengjaroensukying	ID. 46010276
	Mr. Thanapat Wanachart	ID. 46010278
Advisor	Mr. Puchong Hongsuwan	
Graduate Level	Bachelor Degree of Information Engineering	
Department	Information Engineering	
Academic Year	2006	

ABSTRACT

Nowadays, multimedia's become to be a part of present life. This program is created to serve multimedia in part of movie by using subtitle. Program will support to serve for edit copy and modify subtitle that include in movie file. It can detect subtitle in movie and then change it in to text file by use method of Image Processing with accuracy in symbol and time of subtitle. And then it'll be modified by create subtitle file .ass (Advanced Substation Alpha) that can use with another Subtitle Studio.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

ปริญญานิพนธ์ฉบับนี้สำเร็จได้ด้วยดี ด้วยความร่วมมือกันหลายฝ่าย บุคคลแรกที่สำคัญที่สุดที่ต้องขอบคุณ เพราะท่านทำให้เรามีวันนี้ได้ คือ บิดาและมารดาที่เคารพรัก ที่คอยเป็นกำลังสนับสนุน ในด้านการศึกษา และเป็นกำลังมาตลอด

ขอบคุณ อ.ภุชงค์ หงษ์สุวรรณ อาจารย์ที่ปรึกษาปริญญานิพนธ์และคอยช่วยเหลือ ให้คำแนะนำ คำปรึกษา ตรวจทานแก้ไข และให้ข้อคิด รวมทั้งแนวทางในการดำเนินงาน จึงทำให้การจัดทำปริญญานิพนธ์จนสำเร็จลุล่วงด้วยดี ขอขอบพระคุณเป็นอย่างยิ่ง

ขอขอบคุณ ภาควิชาวิศวกรรมสารสนเทศ คณะวิศวกรรมศาสตร์ สำหรับสิ่งอำนวยความสะดวกมากมายที่ใช้ทำการวิจัย ทั้งห้องทำงาน โต๊ะทำงาน อินเทอร์เน็ต สำหรับทำการค้นคว้าหาข้อมูลต่างๆ ที่จำเป็นสำหรับการทำปริญญานิพนธ์ฉบับนี้

รวมทั้งขอบคุณเพื่อนๆ ทุกคน ที่เป็นกำลังใจและ ให้ความร่วมมือทั้งความรู้ที่ได้รับแลกเปลี่ยนแนวคิดการทำงาน และเพื่อนผู้ซึ่งร่วมผ่าฟันไปด้วยกัน

นางสาว ณิชจิรา ชอบธรรม

นาย ธนพล จึงเจริญสุขยิ่ง

นาย ธนพัฒน์ วรรณชาติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

	หน้า
บทคัดย่อ ภาษาไทย	ก
บทคัดย่อ ภาษอังกฤษ	ข
กิตติกรรมประกาศ	ค
สารบัญ	ง
สารบัญรูป	ช
บทที่ 1 บทนำ	1
1.1วัตถุประสงค์	1
1.2 ขอบเขตของโครงงาน	1
1.3 สถาปัตยกรรมของระบบ	2
1.4 ขั้นตอนการดำเนินโครงงาน	3
บทที่ 2 ทฤษฎีและหลักการที่นำมาใช้	5
2.1 ไฟล์ข้อมูลภาพยนตร์	5
2.1.1 การบีบอัดไฟล์ภาพยนตร์	5
2.1.1.1 รูปแบบ Format ของไฟล์ภาพยนตร์	5
2.1.1.2 โคเดค (เข้ารหัส /ถอดรหัส)	6
2.1.1.2.1 ดิปเอ็กซ์	6
2.1.1.2.2 เอ็กซ์วิดท์	6
2.1.2 คุณภาพของวิดีโอ	6
2.1.2.1 อัตราเฟรม	6
2.1.2.2 ความละเอียด	7
2.1.3 ไฟล์คำบรรยายใต้ภาพภาพยนตร์	7
2.1.3.1 ประเภทของคำบรรยายข้างใต้ภาพยนตร์	7
2.1.3.2 รูปแบบของไฟล์คำบรรยายข้างใต้ภาพยนตร์	7
2.2 การประมวลผลภาพ	10
2.2.1 การรวมแบนด์	10
2.2.2 ระบบสี อาร์จีบี (RGB)	10
2.2.3 การวิเคราะห์ภาพแบบไบนารี	12

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
2.2.3.1 การพิจารณาว่าพิกเซลเชื่อมต่อแบบใดได้ดังนี้	13
2.2.3.2 การนับจำนวนกลุ่มของพิกเซลที่อยู่ติดกัน	14
2.2.3.3คุณสมบัติทางโทโปโลยีของภาพดิจิตอล	14
2.2.4 การทำเทรสโฮลด์ดิ้ง (Thresholding)	16
2.2.4.1 ค่าเทรสโฮลด์ที่กำหนดล่วงหน้า	17
2.2.4.2 ค่าเทรสโฮลด์ที่กำหนดขึ้นมาจากค่าเฉลี่ย	17
2.2.4.3 ค่าเทรสโฮลด์ที่คำนวณจากวิธีการหาแบบซ้ำๆ	17
2.2.5 การสลับค่า (Invert)	17
2.2.6 การแบ่งออกเป็นส่วน (Segmentation)	18
2.2.6 การทำให้บาง (Thinning)	20
2.3 การจำแนกตัวอักษร (Character Recognition)	21
2.3.1 โอซีอาร์ (OCR หรือ Optical Character Recognition)	23
2.3.2 การรู้จำรูปแบบ (Pattern Recognition)	24
2.3.3 วิธีการวิเคราะห์ทางโครงสร้าง (Structural Analysis)	24
2.3.3.1 การแบ่งภาพอักษรออกเป็น 9 ส่วน	24
2.3.3.2 การวิเคราะห์โดยใช้คุณสมบัติทางโทโพโลจีของภาพ	25
2.3.3.3 การวิเคราะห์หาค่ารหัส 4 บิตลงในภาพอักษร	27
2.3.3.4 การจัดเก็บลงฐานข้อมูลของการรู้จำ	28
2.4 เครื่องมือที่ใช้พัฒนาและไลบรารี่ที่เกี่ยวข้อง	29
2.4.1 การประมวลผลภาพขั้นสูงของจาวา (Java Advance Imaging)	29
2.4.1.1 ขั้นตอนการดำเนินงานและตัวดำเนินการ เจเอไอ	30
2.4.1.2 ตัวดำเนินการของเจเอไอ	32
2.4.1.3 การเข้าถึงข้อมูลของภาพ (Image data access)	34
2.4.2 การจัดการด้านมัลติมีเดียของจาวา (JMF หรือ Java Media Framework)	34
2.4.2.1 การควบคุมตำแหน่งเฟรม (FramePositioningControl)	35
2.4.2.2 การดักจับเฟรมจากวีดีโอสตรีม (FrameGrabbingControl)	37
2.4.2.3 การเปลี่ยนบัฟเฟอร์ไปเป็นอิมเมจ (BufferToImage)	37

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
บทที่ 3 การออกแบบโครงงาน	39
บทที่ 4 ผลการทดลอง	44
4.1 การทดลองการตรวจสอบคำบรรยายภาพยนตร์	44
4.2 ผลการทดลองที่ได้กับโปรแกรมซับไทเทิลสตูดิโอทั่วไป	49
บทที่ 5 สรุปผลการดำเนินงาน	50
5.1 ปัญหาที่พบในระหว่างการดำเนินโครงงาน	50
5.2 แนวทางการแก้ไข	50
5.3 แนวทางในการพัฒนา	51
บรรณานุกรม	52

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

	หน้า
รูปที่ 1.1 แผนผังการทำงานของโปรแกรม MSCR	2
รูปที่ 2.1 ภาพที่เกิดจากพิกเซลวางเรียงต่อกัน	10
รูปที่ 2.2 ภาพระบบสี อาร์จีบี	11
รูปที่ 2.3 ภาพภาพประกอบของภาพ อาร์จีบี	11
รูปที่ 2.4 ภาพตัวอย่างระดับสี อาร์จีบี ในระบบ 24 บิท	12
รูปที่ 2.5 ตัวอย่างพิกเซลที่เรียงกันเป็นตัวอักษร	12
รูปที่ 2.6 ลักษณะของพิกเซลเพื่อนบ้าน	13
รูปที่ 2.7 ลักษณะของพิกเซลเพื่อนบ้าน (2)	14
รูปที่ 2.8 หน้าต่างขนาด 3 x 3 บิต	14
รูปที่ 2.9 ลักษณะคุณสมบัติทางโทโปโลจิคอลของจุดภายใน	15
รูปที่ 2.10 ลักษณะคุณสมบัติทางโทโปโลจิคอลของจุดโดดเดี่ยว	15
รูปที่ 2.11 ลักษณะคุณสมบัติทางโทโปโลจิคอลของจุดตัด	15
รูปที่ 2.12 ลักษณะคุณสมบัติทางโทโปโลจิคอลของจุดต่อ	15
รูปที่ 2.13 ลักษณะคุณสมบัติทางโทโปโลจิคอลของจุดแยก	16
รูปที่ 2.14 ลักษณะคุณสมบัติทางโทโปโลจิคอลของจุดปลาย	16
รูปที่ 2.15 ภาพการทำสลับค่า	18
รูปที่ 2.16 Line Detection	19
รูปที่ 2.17 Gap Detection	19
รูปที่ 2.18 Space Detection	20
รูปที่ 2.19 รูปตัวอักษรที่ทำการแบ่งออกเป็นส่วนจากภาพ	20
รูปที่ 2.20 ภาพการทำ Thinning	21
รูปที่ 2.21 โครงสร้างทั่วไปของการจำแนกตัวอักษร	21
รูปที่ 2.22 รูปตัวอักษร “A” ที่นำมาคำหาโทโพโลจี	25
รูปที่ 2.23 ตัวอย่างของจุดภาพที่จะนำมาใช้วิเคราะห์	25
รูปที่ 2.24 ตัวอย่างลักษณะของคุณสมบัติโทโปโลยีในภาพอักษร	27
รูปที่ 2.25 ค่ารหัส 4 บิต จุดแยก, จุดปลาย, จุดตัด, จุดต่อ	28

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

	หน้า
รูปที่ 2.26 ค่ารหัส 4 บิต ในจุดทั้ง 9 ของภาพตัวอักษร “A”	28
รูปที่ 2.27 ค่าประจำหลักของแต่ละจุดในค่ารหัส 4 บิต	29
รูปที่ 2.28 ค่ารหัส 4 บิตลงในภาพอักษรหลังจากคูณด้วยค่าประจำตำแหน่ง	29
รูปที่ 2.29 โครงสร้างของพลานาอิมเมจ	31
รูปที่ 2.30 ดูไฟล์ภาพยนตร์โดยเจเอมสตูดีโอ (JMStudio) ประยุกต์จาก เจเอมเอฟ	35
รูปที่ 3.1 Flow-chart การทำงานของขั้นตอนการทำงานหลัก	39
รูปที่ 3.2 Flow-chart การทำงานของ Image Processing	40
รูปที่ 3.3 Flow-chart การทำงานของ Threshold	41
รูปที่ 3.4 Flow-chart การ Segment แต่ละตัวอักษร	42
รูปที่ 3.5 Flow-chart การคิดค่าTopology ของแต่ละตัวอักษรในกระบวนการ Recognition	43
รูปที่ 4.1 หน้าจอการเริ่มต้นของโปรแกรม	44
รูปที่ 4.2 หน้าจอการเลือกไฟล์เพื่อไปกำหนดเป็น Data Source ของโปรแกรม	45
รูปที่ 4.3 หน้าจอการทำงานของโปรแกรม	46
รูปที่ 4.4 หน้าจอการผลการทำงานโดยการอัพเดทซับไตเติ้ลไลน์	47
รูปที่ 4.5 หน้าจอการบันทึก Subtitle line ในนามสกุล .เอเอสเอสใน Directory ที่ต้องการ	48
รูปที่ 4.6 หน้าจอผลที่ได้โดยโปรแกรม ซับไทเทิลสตูดิโอ Aegisub v1.10	49

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

ปัจจุบันการนำเสนอและการบริการด้านมัลติมีเดียเป็นส่วนสำคัญส่วนหนึ่งในการดำเนินชีวิตในสังคมปัจจุบัน ข้อมูลข่าวสารตลอดไปจนถึงการด้านความบันเทิง ล้วนมีการนำเทคโนโลยีด้านมัลติมีเดียเข้ามามีส่วนร่วมในการนำเสนอ เพื่อดึงดูดความสนใจของผู้รับชมข่าวสาร

การเสนอการบริการหรือรูปแบบการทำงานในส่วนต่างๆ มีการแข่งขันกันอย่างแพร่หลายในด้านความแตกต่าง รูปแบบ ความรวดเร็ว ซึ่งต้องออกแบบมาให้ตอบสนองกับความต้องการของผู้บริโภค จึงมีการพัฒนา รูปแบบของ ตัวโปรแกรมที่นำเสนอ รูปแบบการทำงานที่แตกต่าง แนวคิดหรือการออกแบบที่ไม่ซ้ำกับแบบเดิมที่มีอยู่

โดยโปรแกรมนี้ออกแบบมาให้ผู้ใช้สามารถ สร้างคำบรรยายภาพยนตร์ อย่างง่ายด้วยตัวเอง โดยอาศัยหลักการของการจดจำตัวอักษร (Character Recognition) เพื่อนำไปใช้ในการแก้ไข และแสดงคำบรรยายในรูปแบบอื่นๆ ต่อไป

1.1 วัตถุประสงค์

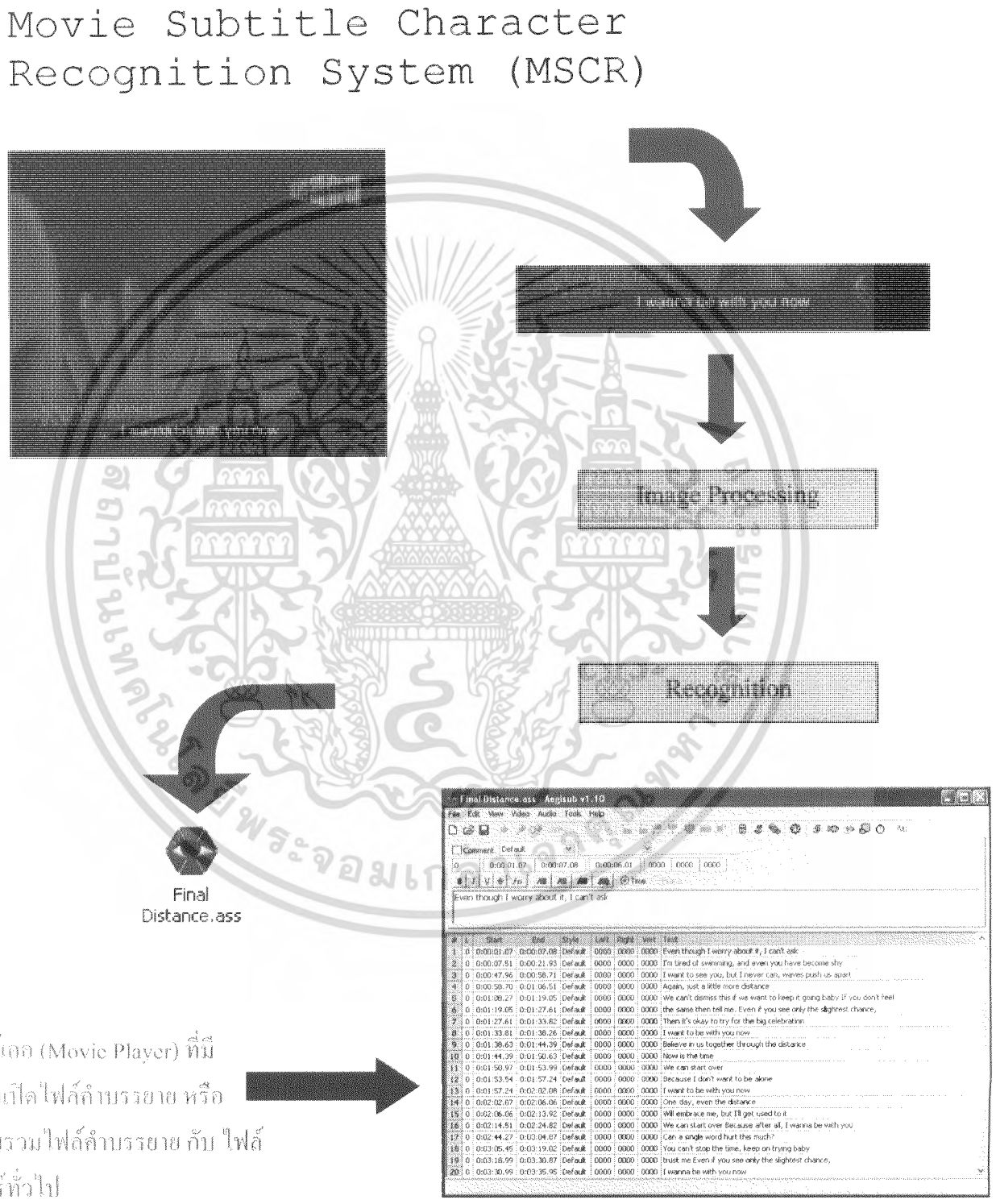
- เพื่อให้ผู้ใช้งานสามารถแก้ไขคัดลอกคำบรรยายภาพยนตร์ที่มีการเก็บข้อมูลข้อเวลาเริ่มต้นแสดงคำบรรยาย เวลาสิ้นสุดของคำบรรยายและตัวอักษรที่อยู่ในคำบรรยายว่าเป็นตัวอักษรใด
- เพื่อที่สามารถทำคำบรรยายภาพยนตร์ได้สะดวกยิ่งขึ้นโดยไม่จำเป็นต้องทำการ timing เอง โดยอาศัยการดึงจากคำบรรยายจากคำบรรยายภาพยนตร์ที่มีอยู่แล้ว

1.2 ขอบเขตของโครงงาน

- สามารถจดจำตัวอักษรจากภาพยนตร์ได้
- มีความถูกต้องของตัวอักษรที่จดจำได้ ทั้งด้าน เวลา ตัวอักษร
- สามารถชมภาพตัวอย่างพร้อมๆกับการทำงานของโปรแกรมได้
- สามารถนำผลที่ได้มาแสดงในรูปแบบไฟล์ คำบรรยาย เช่น .เอสเอสเอ (.ssa) , .เอเอสเอส (.ass), .ซับ(.sub) เนื่องจากรูปแบบไฟล์เหล่านี้เป็นที่นิยม และสามารถใช้ร่วมกับโปรแกรมในการนำเสนอต่างๆ เช่น พวก ไดเรคท์วอปซับ (DirectVobSub) หรือ โปรแกรมในการทำฮาร์ดซับ (Hard Sub) อื่นๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.3 สถาปัตยกรรมของระบบ



รูปที่ 1.1 แผนผังการทำงานของโปรแกรม MSCR

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สถาปัตยกรรมของระบบประกอบไปด้วย 2 ส่วนหลักคือ

การประมวลผลภาพ (Image Processing) คือ ขั้นตอนในการเตรียมภาพตัวอักษร ก่อนที่จะนำไปวิเคราะห์รู้จำตัวอักษร เช่น การปรับภาพให้เป็นภาพไบนารี, การแยกตัวอักษรแต่ละตัวออกจากข้อความ, การทำตัวอักษรให้บาง

การรู้จำตัวอักษร (Recognition) คือ การวิเคราะห์จากโครงสร้างของตัวอักษรโดยจะ แบ่งตัวอักษรที่ผ่านขั้นตอนการเตรียมตัวอักษรมาแล้ว ออกเป็น 9 ส่วน แล้วพิจารณาเปรียบเทียบหารหัสที่ใช้แทนตัวอักษร โดยการวิเคราะห์จากคุณสมบัติทางโทโปโลยีของจุดภาพในแต่ละส่วน
การทำงานของโปรแกรมต้องมีส่วนประกอบเริ่มต้นในการทำงานคือ

1. ภาพยนตร์ที่มีการใส่คำบรรยายไว้แล้ว
2. ตัวโปรแกรม Movie Subtitle Character Recognition (MSCR)

โดยการทำงานจะเริ่มจากการนำไฟล์ภาพยนตร์ มาแสดงผลกับโปรแกรม MSCR เพื่อให้ได้ไฟล์คำบรรยายภาพยนตร์ซึ่งเป็นประเภทซอฟท์ซับ (Soft Sub) ออกมา เพื่อที่จะสามารถนำไปใช้ในการทำงานอื่นๆ ต่อไป

1.4 ขั้นตอนการดำเนินโครงงาน

1. ดำเนินการค้นคว้าหา วิธีการ เทคนิค และตัวดำเนินการที่เหมาะสม (โปรแกรมที่ใช้)

ในขั้นต้นของโครงงานเป็นการศึกษาหาข้อมูลแนวทางในการดำเนินการ ในการออกแบบโดยหาข้อมูลว่าโปรแกรมใดเหมาะสมในการทำงานในขอบเขตที่กำหนด โดยพิจารณาว่า รูปแบบการทำงานของโปรแกรมเอื้ออำนวยต่อการทำงานหรือไม่ มีโครงการที่คล้ายคลึงกันที่ใช้โปรแกรมนี้ในการทำงานหรือไม่

2. วิเคราะห์ข้อมูลที่ได้เพื่อหาความเป็นไปได้ในการออกแบบ

ทำการศึกษาข้อมูลว่าการออกแบบโปรแกรมนั้นมีความเป็นไปได้ในการใช้งานจริงหรือมีโอกาส ที่จะประสบผลสำเร็จจริงหรือไม่ ทำอย่างไรจริงจะสอดคล้องกันระหว่างผลที่ต้องการและโอกาสสำเร็จในการทำงาน

3. เขียนโปรแกรมและดำเนินการทดสอบ

เขียนโปรแกรมตามแผนงานที่วางไว้ และทดสอบดูเป็นระยะๆว่าสามารถที่จะดำเนินการตามแผนที่วางไว้ได้หรือไม่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. ปรับปรุงแก้ไขโปรแกรมที่ได้หรือเปลี่ยนวิธีการดำเนินงาน

ทำการแก้ไขรูปแบบของโปรแกรมหรือเปลี่ยนวิธีในการทำ เช่น การเปลี่ยนลำดับขั้นตอนของการทำ การประมวลผลภาพ(Image Processing)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
ทฤษฎีทั่วไป

ระบบจดจำตัวอักษรคำบรรยายภาพยนตร์ คือ ระบบที่นำไฟล์ข้อมูลภาพยนตร์มาทำการประมวลผลภาพ (Image Processing) เพื่อให้ได้คำบรรยายข้างใต้ภาพยนตร์ (Subtitle) แล้วนำมาทำการจำแนกตัวอักษร (Character Recognition) เป็นตัวอักษรซึ่งจะถูกบันทึกเป็นไฟล์คำบรรยายข้างใต้ภาพยนตร์ เช่น *.เอสเอสเอ (ssa) , *.เอเอสเอส (ass) , *.ซับ (sub) เพื่อนำไปใช้ต่อไป

การศึกษาและเข้าใจในกระบวนการของระบบดังที่กล่าวมาข้างต้นสามารถแยกเป็นหัวข้อได้ดังนี้

2.1 ไฟล์ข้อมูลภาพยนตร์

2.2 การประมวลผลภาพ

2.3 การจำแนกตัวอักษร

2.4 เครื่องมือที่ใช้พัฒนาและไลบรารี่ที่เกี่ยวข้อง

2.1 ไฟล์ข้อมูลภาพยนตร์

2.1.1 การบีบอัดไฟล์ภาพยนตร์

การบีบอัดไฟล์ข้อมูลเป็นการขนาดของไฟล์มีขนาดเล็กลง แต่ยังคงมีปริมาณของข้อมูลเท่าเดิม เช่นก่อนผ่านกระบวนการบีบอัดข้อมูล หรือที่เรียกว่า “โคเดค (Codec)” ซึ่งทำให้ไฟล์ที่มีขนาด 100 เมกะไบต์ (MB) หลังจากที่ทำการบีบอัดแล้ว ขนาดของไฟล์อาจลดลงเหลือเพียง 11 เมกะไบต์ โดยที่ไฟล์อาจจะมีคุณภาพภาพและเสียงลดลง ขึ้นอยู่กับรูปแบบการบีบอัดข้อมูลซึ่งยังคงมีการพัฒนาต่อมาเรื่อยๆ สามารถแสดงได้ดังต่อไปนี้

2.1.1.1 รูปแบบ Format ของไฟล์ภาพยนตร์

เอวีไอ (Audio Video Interleave) เป็นรูปแบบแฟ้มมัลติมีเดียสำหรับใช้กับวินโด้มีเดียเพลเออ (Windows Media Player) รูปแบบแฟ้มนี้สามารถมีได้ทั้งเสียงและวิดีโอ และรูปแบบแฟ้มนี้ยึดตามข้อกำหนดของ มาตรฐานข้อตกลงการแลกเปลี่ยนรูปแบบไฟล์ (RIFF) รูปแบบแฟ้มเอวีไอจะมีนามสกุลของชื่อแฟ้มเป็น .เอวีไอ ถูกพัฒนาขึ้นโดยบริษัทไมโครซอฟท์ (Microsoft) ในปี 1992 บนระบบปฏิบัติการวินโดว์ (Windows) ภายในไฟล์เอวีไอประกอบด้วยข้อมูลประเภทภาพและเสียง สนับสนุนการเล่นแบบย้อนกลับและสนับสนุน มัลติเพิลออดิโอ (Multiple Audio) และวีดีโอสตรีม (Video Streams) ข้อเสียคือมีขนาดใหญ่ กินพื้นที่ขนาดบน ฮาร์ดดิสก์(Hard disk) มาก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.1.2 โคเดค (เข้ารหัส /ถอดรหัส)

เป็นอัลกอลิทึมที่ใช้ในการแปลงข้อมูลภาพและเสียงระหว่างอะนาล็อกกับดิจิตอล (หรือระหว่างดิจิตอลด้วยกัน) ตัวอย่างเช่น ดิปเอ็กซ์และเอ็กซ์วิดท์ (XVid) ทั้ง 2 แบบนี้จะใช้สำหรับวิดีโอดิจิตอลที่มีคุณภาพสูงๆ ส่วนใหญ่แล้ว โคเดคจะสามารถสร้างไฟล์ในรูปแบบ เอวีไอ และสามารถใช้ร่วมกันกับเครื่องพีซีได้ด้วย

2.1.1.2.1 ดิปเอ็กซ์ ผู้พัฒนา MPEG-4 part 2 รายแรกๆ คือไมโครซอฟท์ (.asf) และดิปเอ็กซ์ในยุคแรกๆ ก็เป็นเวอร์ชันที่แฮค .asf ให้เก็บเป็น .เอวีไอ ได้ในภายหลังซึ่งต่อมาดิปเอ็กซ์ ได้แก้ไขให้เป็นอัลกอริทึมของตัวเองและแจกให้ใช้ฟรี ส่วน source นั้นเป็นกรรมสิทธิ์ของบริษัทดิปเอ็กซ์ Networks ซึ่งก่อตั้งโดยบรรดาแฮคเกอร์ที่แฮคไมโครซอพท์ ปัจจุบัน ดิปเอ็กซ์ ได้รับความนิยมสูงมากโดยเฉพาะโลกของหนังที่เข้ารหัสใหม่ (.rip) มาจากดีวีดี และอนิเมแฟนซับ เพราะได้คุณภาพเท่าดีวีดีในขนาดเท่าซีดี

ดิปเอ็กซ์ สามารถลดข้อมูลเหลือเพียง 10 –20 เปอร์เซ็นต์ของปริมาณข้อมูลเดิม เช่น สามารถบีบอัดสัญญาณจาก ดีวีดี ลดลงจากปกติ 75-80% คือเหลือประมาณประมาณ 650 MB ที่ความละเอียด 640 X 480 พิกเซล (pixels) หรือย่อภาพยนตร์ยาวๆ หนึ่งเรื่องมาเก็บในแผ่น CD –R เพียงแผ่นเดียว และยังสามารถเปิดชมภาพยนตร์ด้วยโปรแกรมธรรมดา เช่น มีเดียเพลเออ ได้อีกด้วย

2.1.1.2.2 เอ็กซ์วิดท์ เมื่อ ดิปเอ็กซ์Networks ได้รับความนิยม ก็โอเพ่นซอร์ข้อมูลบางส่วนใต้โครงการโอเพนดิปเอ็กซ์ (OpenDivX) และมีนักพัฒนาสนใจใช้งานกันเยอะ แต่ภายหลังดิปเอ็กซ์ Networks ได้เปลี่ยนใจหยุดพัฒนาโครงการโอเพนดิปเอ็กซ์ ทางชุมชนจึงแก้ปัญหาโดยการพัฒนารูปแบบการเข้า/ถอดรหัส (โคเดค) ที่เป็นโอเพ่นซอร์สขึ้นมาแข่งกับ ดิปเอ็กซ์ และใช้ชื่อชนกันว่า เอ็กซ์วิดท์ (เขียนกลับหลัง) ระดับความนิยมนั้นใกล้เคียงกับ ดิปเอ็กซ์

2.1.2 คุณภาพของวิดีโอ

การวัดคุณภาพของวิดีโอสามารถวัดได้จาก อัตราเฟรม (Frame Rate) และความละเอียด (Resolution) ของภาพ (ปัจจุบันอาจนำขนาดของข้อมูลหลังการบีบอัดมาพิจารณาด้วยกรณีที่ทั้งอัตราเฟรมและความละเอียดเท่ากัน)

2.1.2.1 อัตราเฟรม

อัตราเฟรม คือ อัตราความถี่ในการแสดงภาพจาก Timeline ออกทางหน้าจอ อัตราที่เฟรมถูกแสดงในไฟล์ภาพยนตร์มีหน่วยเป็นเฟรมต่อวินาที แสดงในรูปของ FPS ซึ่งย่อมาจาก อัตราเฟรมต่อวินาที (Frame Per Second) เป็นหน่วยวัดปริมาณข้อมูลที่ใช้ในการเก็บบันทึกและแสดงวิดีโอ นอกจากนี้ยังสามารถแสดงเฟรมให้มีความต่อเนื่องในเวลาอันรวดเร็ว)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.2.2 ความละเอียด

ความละเอียด หมายถึง ความคมชัดของภาพที่แสดงออกทางจอภาพ โดยความละเอียดของจอภาพขึ้นอยู่กับจำนวนจุดทั้งหมดที่เกิดบนจอ จุดต่างๆ นี้เรียกว่า พิกเซล

2.1.3 ไฟล์คำบรรยายใต้ภาพภาพยนตร์

2.1.3.1 ประเภทของคำบรรยายข้างใต้ภาพยนตร์

1) ฮาร์ดซับ (Hard sub)

คือ จะมีการใส่ คำบรรยายข้างใต้ภาพยนตร์ ฝังลงไปในตัวภาพยนตร์เลย ดังนั้นในการรับชมจะไม่สามารถเลือก ปิด คำบรรยายข้างใต้ภาพยนตร์ ได้ และ คำบรรยายข้างใต้ภาพยนตร์จะมีเพียง 1 ภาษาเท่านั้น

ข้อดีของฮาร์ดซับ คือ ตัวหนังสือจะมีขนาดใหญ่พอดี และ แสดงเป็นภาษาไทยได้แน่นอน เพราะมีการเซ็ทค่ามาแล้ว และใส่รวมลงไปกับไฟล์ภาพยนตร์ จึงใช้กับเครื่องเล่นได้ทุกรุ่น

2) ซอฟซับ (Soft sub)

โดยทั่วไป ดิปเอ็กซ์ (DivX) และ เอ็กซ์วิดท์ จะเป็นแบบ ซอฟซับ คือใช้ไฟล์ คำบรรยายข้างใต้ภาพยนตร์ (.เอสอาร์ที (.str) หรือ .ซับ (.sub)) แยกออกจากตัวไฟล์ภาพยนตร์ซึ่งต้องการ software ในการที่ทำหน้าที่ซ้อนคำบรรยายใต้ภาพภาพยนตร์ลงบนวิดีโอ ในการเลือกแสดงคำบรรยายข้างใต้ภาพยนตร์ ปรับขนาดตัวหนังสือ หรือจัดวางตำแหน่งของคำบรรยายข้างใต้ภาพยนตร์ซึ่งไฟล์คำบรรยายข้างใต้ภาพยนตร์จะต้องอยู่ในโฟลเดอร์เดียวกับไฟล์ภาพยนตร์ ในส่วนของการแสดงผล เช่น ขนาดของตัวหนังสือ, สีของตัวหนังสือ, ตำแหน่ง และ ภาษาที่สนับสนุน จะขึ้นอยู่กับความสามารถของเครื่องเล่นแต่ละยี่ห้อ ซึ่งมีไม่เท่ากัน บางรุ่นอาจจะแสดงตัวหนังสือเล็กมาก หรือ ไม่รองรับภาษาไทย

ข้อดีของซอฟซับ คือ สามารถเลือกชม คำบรรยายข้างใต้ภาพยนตร์ ได้หลายภาษา เช่น ไทย, อังกฤษ, จีน, เกาหลี ซึ่งต้องใช้กับเครื่องเล่น ดีวีดี (DVD) สนับสนุนภาษาต่างๆเหล่านี้ ได้ และครื่องบางรุ่นสามารถ การยกระดับเฟิร์มแวร์ (firmware) เพื่อให้มีความสามารถเพิ่มเติมในการปรับแต่งลูกเล่นเกี่ยวกับ คำบรรยายข้างใต้ภาพยนตร์ หรือภาษาที่รองรับได้

2.1.3.2 รูปแบบของไฟล์คำบรรยายข้างใต้ภาพยนตร์

1) เอสเอสเอ (SSA) หรือซับสเตชั่นอัลฟา (Sub Station Alpha)

เป็นรูปแบบของคำอธิบายหรือคำแปลที่มักจะอยู่ข้างใต้ของภาพยนตร์ซึ่งเป็นที่นิยมกันอย่างแพร่หลายของผู้เรียบเรียงคำแปล ถูกพัฒนาหลัง ไฟล์ แบบเอสทีอาร์ รองรับการทำงานของ การกำหนดตำแหน่ง, คาราโอเกะ (karaoke), การจัดการรูปแบบ ฯลฯ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปแบบมาตรฐานไฟล์เอสเอสเอ

Title: Neon Genesis Evangelion - Episode 26

Original Script: RoRo

Script Updated By: version 2.8.01

ScriptType: v4.00

Collisions: Normal

PlayResY: 600

PlayDepth: 0

Timer: 100,0000

[V4 Styles]

Format: Name, Fontname, Fontsize, PrimaryColour, SecondaryColour, TertiaryColour, BackColour, Bold, Italic, BorderStyle, Outline, Shadow, Alignment, MarginL, MarginR, MarginV, AlphaLevel, Encoding

Style: DefaultVCD, Arial,28,11861244,11861244,11861244,-2147483640,-1,0,1,1,2,2,30,30,30,0,0

[Events]

Format: Marked, Start, End, Style, Name, MarginL, MarginR, MarginV, Effect, Text

Dialogue: Marked=0,0:00:01.18,0:00:06.85,DefaultVCD, NTP,0000,0000,0000,,{\pos(400,570)}Like an angel with pity on nobody

2) เอเอสเอส (.ASS) หรือแอดวานซ์ซับสเตชั่นอัลฟา (Advanced SubStation Alpha)

เป็นเทคนิคแบบ เอสเอสเอ เวอร์ชั่น4+ มีรูปแบบคล้ายคลึงกับ .เอสเอสเอ มาก สามารถสร้าง คำบรรยายข้างใต้ภาพยนตร์ที่มีประสิทธิภาพมากกว่าแบบเอสเอสเอ เช่นสร้าง คำบรรยายข้างใต้ภาพยนตร์แบบแมนนวลกราฟฟิค (manual graphic) ที่ใช้ในคาราโอเกะได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปแบบมาตรฐานไฟล์เอเอสเอส

Title: Neon Genesis Evangelion - Episode 26

Original Script: RoRo

Script Updated By: version 2.8.01

ScriptType: v4.00+

Collisions: Normal

PlayResY: 600

PlayDepth: 0

Timer: 100,0000

Video Aspect Ratio: 0

Video Zoom: 6

Video Position: 0

[V4+ Styles]

Format: Name, Fontname, Fontsize, PrimaryColour, SecondaryColour, OutlineColour, BackColour, Bold, Italic, Underline, StrikeOut, ScaleX, ScaleY, Spacing, Angle, BorderStyle, Outline, Shadow, Alignment, MarginL, MarginR, MarginV, Encoding

Style: DefaultVCD, Arial,28,&H00B4FCFC,&H00B4FCFC,&H00000008,&H80000008,-1,0,0,0,100,100,0.00,0.00,1,1.00,2.00,2,30,30,30,0

[Events]

Format: Layer, Start, End, Style, Name, MarginL, MarginR, MarginV, Effect, Text

Dialogue: 0,0:00:01.18,0:00:06.85,DefaultVCD, NTP,0000,0000,0000,,{\pos(400,570)}Like an Angel with pity on nobody

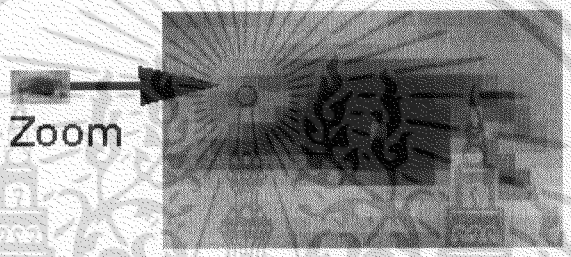
โดยเอเอสเอสนั้นเป็นรูปแบบของวอปซับ (VobSub) หนึ่งซึ่งสนับสนุนการนับเฟรม (framecount) และจำนวนแถวมากๆ (multiline) นอกจากนี้ยังมี .เอสอาร์ที (.srt) และ .ซับ (.sub) ฯลฯ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2 การประมวลผลภาพ

การประมวลผลภาพ คือการเอาภาพมาประมวลผลด้วยคอมพิวเตอร์ เพื่อให้คอมพิวเตอร์รู้ว่าภาพนั้นคือภาพอะไร หรือมีสิ่งที่สนใจอยู่ในภาพหรือไม่ โดยที่ไม่ต้องใช้สายตาของมนุษย์ในการตัดสิน

ภายในภาพเหล่านั้นประกอบด้วยเม็ดสีมากมายที่วางเรียงกันอย่างเป็นลำดับที่แน่นอน เม็ดสีที่เล็กที่สุดเราเรียกว่า พิกเซล (พิกเซล) เมื่อขยายบริเวณดังกล่าวจะพบว่าประกอบด้วยพิกเซลต่าง ๆ จำนวนมากวางเรียงต่อกันไป



รูปที่ 2.1 ภาพที่เกิดจากพิกเซลวางเรียงต่อกัน

2.2.1 การรวมแบนด์

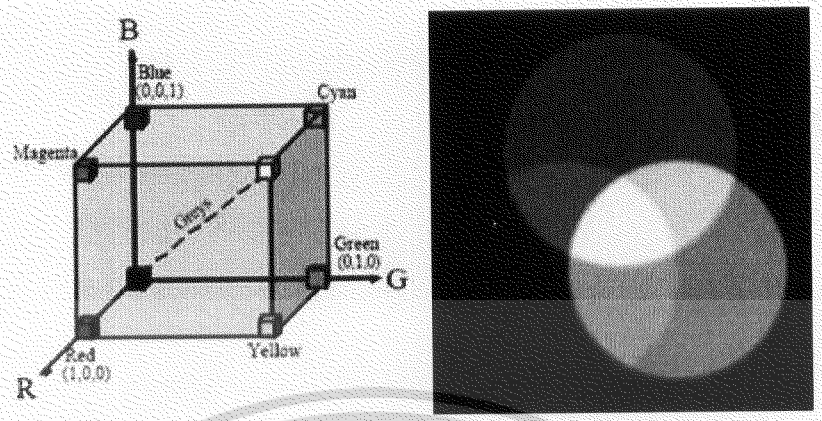
เป็นการรวมกันเพื่อหาค่าเฉลี่ยของภาพในแต่ละแบนด์ โดยทำการรวมทุกแบนด์เข้าด้วยกันเป็นแบนด์เดียวเพื่อที่จะได้สามารถนำไปวิเคราะห์หรือนำไปประมวลผลภาพต่อไป

ในการใช้งานจะใช้ภาพที่ได้จากไฟล์ภาพยนตร์ซึ่งตรงกับมาตรฐาน JPEG จะประกอบด้วย 3 แบนด์ คือ R, G, B จะใช้การรวมแบนด์ ทั้ง 3 ให้เป็นแบนด์ เดียว

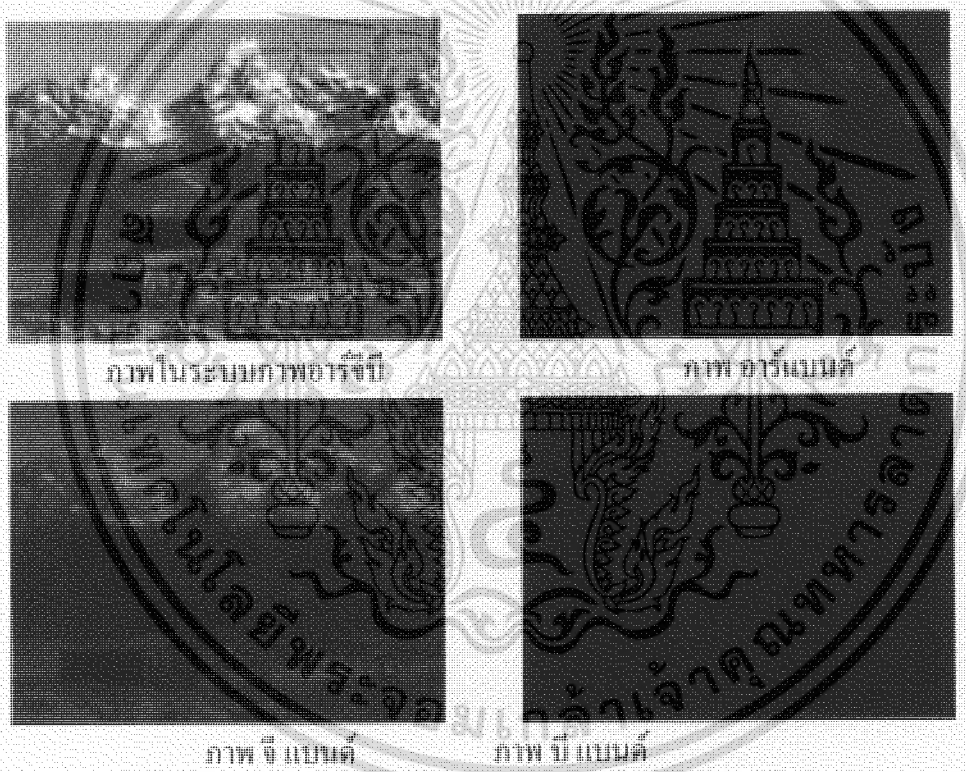
2.2.2 ระบบสีอาร์จีบี (RGB)

ระบบสีอาร์จีบีเป็นระบบสีที่เกิดจากการรวมกันของแสงสีแดง เขียวและน้ำเงินโดยมีการรวมกันแบบแอดดิทีพ (Additive) ซึ่งโดยปกติจะนำไปใช้ในจอภาพแบบซีอาร์ที (CRT) หรือคาโทดเรย์ทิว (Cathode ray tube) ในการใช้งานระบบสีอาร์จีบี ยังมีการสร้างมาตรฐานที่แตกต่างกันออกไปที่นิยมใช้งานได้แต่ อาร์จีบีซีไอเอส (RGBCIE) และอาร์จีบีเอนทีเอสซี (RGBNTSC)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

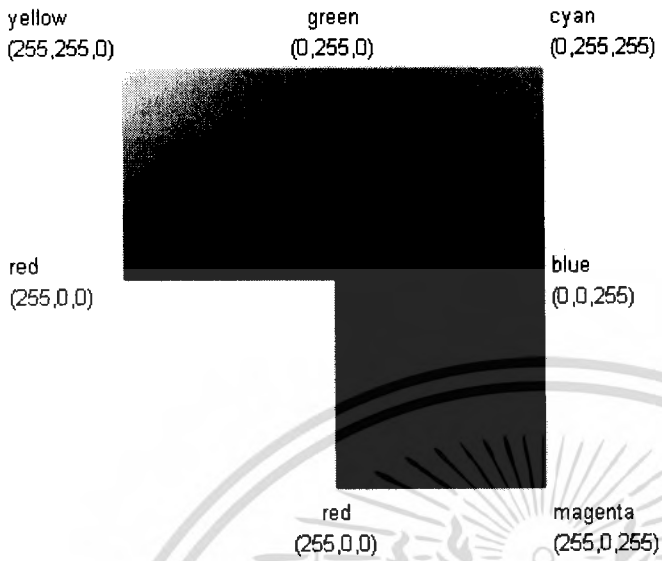


รูปที่ 2.2 ภาพระบบสี อาร์จีบี



รูปที่ 2.3 ภาพประกอบของภาพ อาร์จีบี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(0, 0, 0) เป็นสีดำ

(255, 255, 255) เป็นขาว

(255, 0, 0) เป็นสีแดง

(0, 255, 0) เป็นสีเขียว

(0, 0, 255) เป็นสีน้ำเงิน

(255, 255, 0) เป็นสีเหลือง

(0, 255, 255) เป็นสีเขียวน้ำเงิน

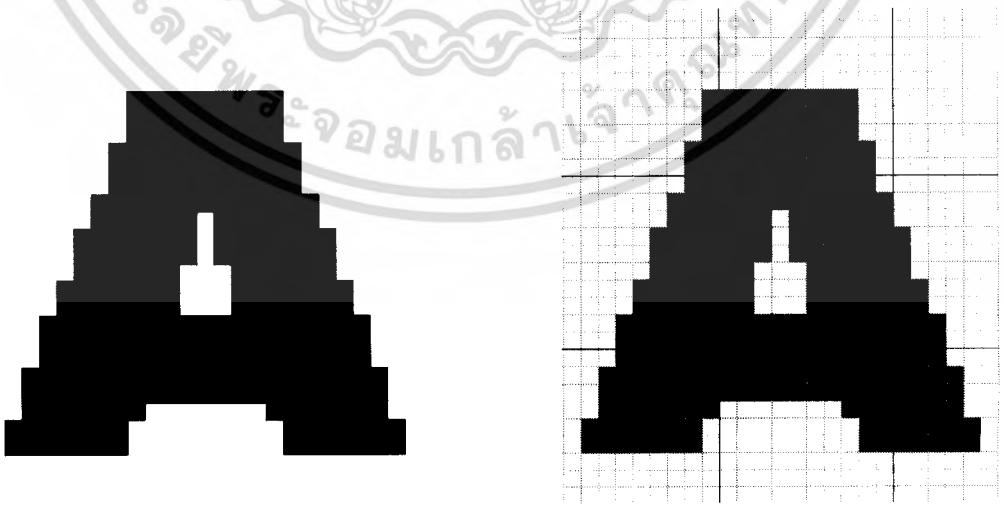
(255, 0, 255) เป็นสีม่วง

รูปที่ 2.4 ภาพตัวอย่างระดับสี อาร์จีบี ในระบบ 24 บิท

2.2.3 การวิเคราะห์ภาพแบบไบนารี

ภาพแบบไบนารี คือ ภาพที่ในแต่ละพิกเซล จะมีความเข้มเพียง 2 ระดับ คือ 0,1 เท่านั้น โดยใช้ 0 หรือ 1 ในการบอกว่าเป็นสีขาวหรือสีดำ การวิเคราะห์ภาพแบบไบนารี ซึ่งจะเลือกสีหนึ่งๆ ในการพิจารณาภาพ โดยจะทำการแบ่งภาพออกเป็นช่องๆ เรียกว่า "พิกเซล"

ในทางปฏิบัติ สิ่งเราที่สนใจ คือกลุ่มของพิกเซลที่อยู่ติดกันหรือที่เรียกว่าพิกเซลเพื่อนบ้าน (connected) เช่นกลุ่ม ของพิกเซลที่เรียงกันเป็นตัวอักษรดังรูปที่ 1



รูปที่ 2.5 ตัวอย่างพิกเซลที่เรียงกันเป็นตัวอักษร

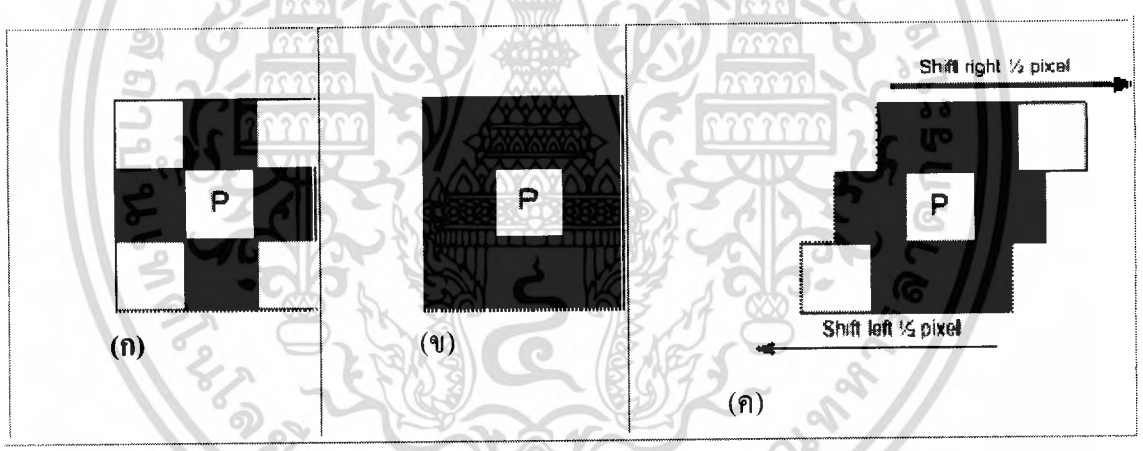
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.3.1 การพิจารณาว่าพิกเซลเชื่อมต่อแบบใดได้ดังนี้

1) แบบสี่เพื่อนบ้าน (4-Connectedness) จะถือว่าแต่ละพิกเซลอยู่ติดกับสี่พิกเซลที่อยู่ด้านบน, ล่าง, ซ้าย และขวา เช่นในรูปที่ 2.6 (ก) เราจะถือว่าพิกเซล P อยู่ติดกับพิกเซลที่เป็นสีแดง ส่วนที่เป็นสีขาวจะไม่ถือว่าเป็นที่ติดกับพิกเซล P

2) แบบแปดเพื่อนบ้าน (8-connectedness) จะถือว่าแต่ละพิกเซลอยู่ติดกับทั้งแปดพิกเซลที่อยู่ล้อมรอบทั้งหมดดังรูปที่ 2.6 (ข)

3) แบบหกเพื่อนบ้าน (6-connectedness) ทั้งสองแบบที่กล่าวมาอาจไม่สอดคล้องบางประการกับเรขาคณิตที่เราคุ้นเคย กล่าวคือจากทฤษฎีเส้นโค้งของจอร์แดน (Jordan's curve theorem) เส้นโค้งบนระนาบที่ลากเป็นวงปิดจะแบ่งระนาบออกเป็นสองส่วนที่ไม่ติดกัน คือส่วนที่อยู่ภายในวงและส่วนที่อยู่นอกวง ดังนั้นวิธีแก้ปัญหาวิธีหนึ่งก็คือการเปลี่ยนไปใช้พิกเซลที่เป็นรูปหกเพื่อนบ้าน โดยกระทำได้โดยกำหนดให้พิกเซลที่อยู่แถวบนเลื่อนไปทางขวา x พิกเซลอยู่แถวล่างเลื่อนไปทางซ้าย x พิกเซล ดังรูปที่ 2.6 (ค)



(ก) สี่เพื่อนบ้าน (ข) แปดเพื่อนบ้าน (ค) หกเพื่อนบ้าน

รูปที่ 2.6 ลักษณะของพิกเซลเพื่อนบ้าน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

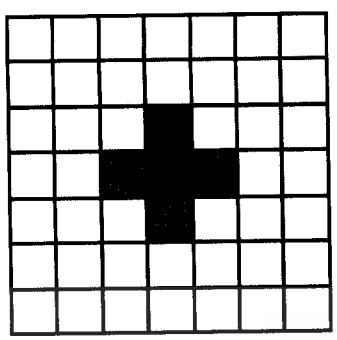


Figure 3a
Rectangular sampling
4-connected

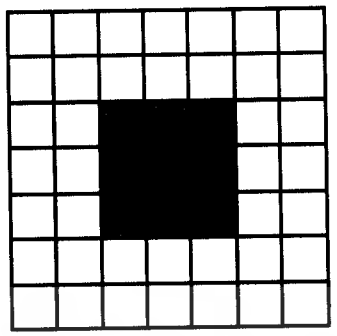


Figure 3b
Rectangular sampling
8-connected

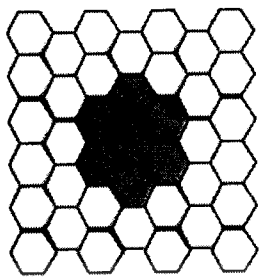


Figure 3c
Hexagonal sampling
6-connected

รูปที่ 2.7 ลักษณะของพิกเซลเพื่อนบ้าน (2)

2.2.3.2 การนับจำนวนกลุ่มของพิกเซลที่อยู่ติดกัน

นิยามกลุ่มของพิกเซลที่อยู่ติดกันในลักษณะคล้ายกับนิยามของ ส่วนประกอบที่อยู่ติดกัน จากทฤษฎีกราฟนั่นคือ ทุกๆ คู่ของพิกเซลในกลุ่มเดียวกันต้องมีเส้นทางเชื่อมไปยังพิกเซลอื่นๆ ในกลุ่มที่เรียงต่อกันโดยต้องไม่มีเส้นทางระหว่างคู่พิกเซลต่างกลุ่ม โดยทำการแจกหมายเลขกลุ่มให้กับแต่ละพิกเซล พิกเซลแต่ละกลุ่มจะได้หมายเลขต่างกัน ภาพไบนารี่ที่นำมาประมวลผลจะถูกเก็บในแถวลำดับ 2 มิติ A[1..n][1..m] ที่มีขนาดเป็น n x m โดยพิกเซลภาพ (หรือวัตถุที่สนใจ) จะมีค่าเป็น 1 และพื้นหลัง (ฉาก) จะมีค่าเป็น 0

2.2.3.3 คุณสมบัติทางโทโปโลยีของภาพดิจิตอล (Topological Properties)

คุณสมบัติทางโทโปโลยีของภาพดิจิตอล จะใช้จำแนกกลุ่มของจุดภาพ โดยวิธีการหาค่าคุณสมบัติทางโทโปโลยีของภาพดิจิตอลนี้ จะใช้วิธีการกำหนดหน้าต่างขนาด 3 x 3 ครอบบิตที่ต้องการหา ซึ่งหน้าต่างขนาด 3 x 3 มีลักษณะดังรูปที่ 2.1 ค่า X0, X1, X2, X3, X4, X5, X6, X7 และX8 เป็นบิตทางดิจิตอลมีค่าเป็น 0 หรือ 1 โดย X0 คือ จุดที่เราจะพิจารณาและจะพิจารณาเฉพาะเมื่อ X0 = 1(จุดภาพเป็นสีดำ) เท่านั้น

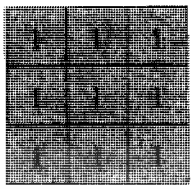
X2	X3	X4
X1	X0	X5
X8	X7	X6

รูปที่ 2.8 หน้าต่างขนาด 3 x 3 บิต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

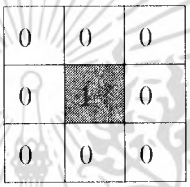
การหาค่าคุณสมบัติทางโทโปโลจิคอล สำหรับภาพดิจิตอลทั่วไป สามารถกำหนดความแตกต่างของกลุ่มจุดภาพในตารางหน้าต่าง 3 x 3 ออกได้เป็น 6 กลุ่ม ดังนี้

1) จุดภายใน (Internal)



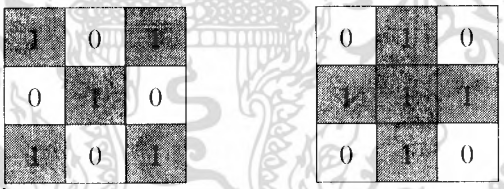
รูปที่ 2.9 ลักษณะคุณสมบัติทางโทโปโลจิคอลของจุดภายใน

2) จุดโดดเดี่ยว (Isolate)



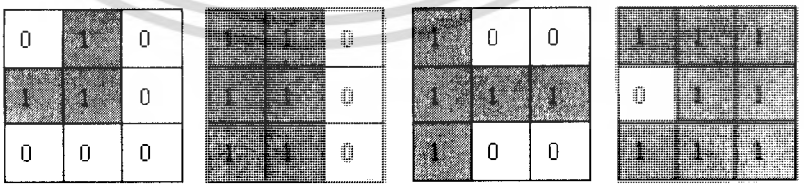
รูปที่ 2.10 ลักษณะคุณสมบัติทางโทโปโลจิคอลของจุดโดดเดี่ยว

3) จุดตัด (Cross)



รูปที่ 2.11 ลักษณะคุณสมบัติทางโทโปโลจิคอลของจุดตัด

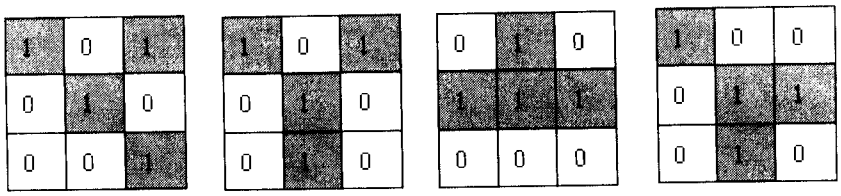
4) จุดต่อ (Connect)



รูปที่ 2.12 ลักษณะคุณสมบัติทางโทโปโลจิคอลของจุดต่อ

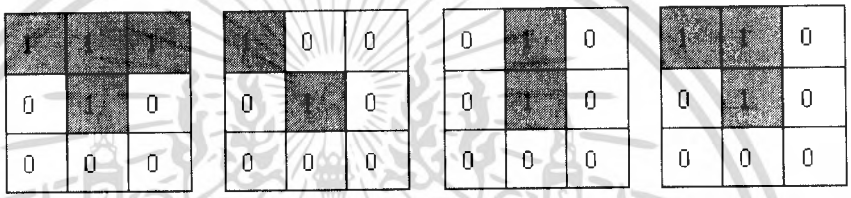
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5) จุดแยก (Branch)



รูปที่ 2.13 ลักษณะคุณสมบัติทางโทโปโลจิคอลของจุดแยก

6) จุดปลาย (End)



รูปที่ 2.14 ลักษณะคุณสมบัติทางโทโปโลจิคอลของจุดปลาย

รูปที่ 2.11 เป็นจุดภายใน จะไม่มีส่วนประกอบรอบๆ ที่มีการเปลี่ยนแปลงจากค่า 1, 0 ไปเป็น 0, 1 เลยและส่วนประกอบรอบๆ จะมีค่าเป็น 1 ทั้งหมด

รูปที่ 2.12 เป็นจุดโดดเดี่ยว จะไม่มีส่วนประกอบรอบๆ ที่มีการเปลี่ยนแปลงจากค่า 0, 1 ไปเป็น 1, 0 เลย และส่วนประกอบรอบๆ จะมีค่าเป็น 0 ทั้งหมด

รูปที่ 2.13 เป็นจุดตัด ที่มีส่วนประกอบรอบๆ ที่มีการเปลี่ยนแปลงจากค่า 1, 0 ไปเป็น 0, 1 อยู่ 8 ช่วง

รูปที่ 2.15 เป็นจุดแยก ที่มีส่วนประกอบรอบๆ ที่มีการเปลี่ยนแปลงจากค่า 1, 0 ไปเป็น 0, 1 อยู่ 6 ช่วง

รูปที่ 2.16 เป็นจุดปลาย ที่มีส่วนประกอบรอบๆ ที่มีการเปลี่ยนแปลงจากค่า 1, 0 ไปเป็น 0, 1 อยู่ 2 ช่วง

2.2.4 การทำเทรสโฮลด์ดิ้ง (Thresholding)

คือการทำการเปลี่ยนภาพระดับเทา (Gray Scale) ให้เป็นภาพขาวดำ (Binary Image) โดยจะใช้ค่าค่าหนึ่งที่กำหนดขึ้นมาเพื่อใช้ตัดสินใจว่าจุดที่มีระดับสีมากกว่าหรือน้อยกว่าเท่าไรจึงจะเป็นสีขาวหรือสีดำ ดังเช่นตามสมการดังนี้

$$g(x,y)=\begin{cases}0 & for\ \ f(x,y)<Threshold\\ 1 & for\ \ f(x,y)\geq Threshold\end{cases} \tag{2.1}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

การหาค่าเทรสโฮลด์สามารถหาได้โดยวิธีต่างๆดังนี้

2.2.4.1 ค่าเทรสโฮลด์ที่กำหนดล่วงหน้า กำหนดค่าเทรสโฮลด์ขึ้นมาตายตัวโดยไม่มีการนำภาพที่ต้องการมาตัดสินใจมาเกี่ยวข้อง

2.2.4.2 ค่าเทรสโฮลด์ที่กำหนดขึ้นมาจากค่าเฉลี่ย (mean) หรือ ค่ากลาง (median) โดยคิดได้จากสมการ

$$Threshold = \frac{\sum_{x=0}^{N-1}\sum_{y=0}^{M-1} f(x,y)}{N * M} \tag{2.2}$$

2.2.4.3 ค่าเทรสโฮลด์ที่คำนวณจากวิธีการหาแบบซ้ำๆ (iterative method) โดยสามารถหาได้ตามขั้นตอนดังนี้

- เลือกค่าเทรสโฮลด์เริ่มต้นด้วยอาจจะสุ่มขึ้นมาหรือจะคิดจากค่าเฉลี่ยก็ได้
- ภาพจะแบ่งเป็นสองส่วนตามที่ได้ตัดเทรสโฮลด์ไว้ในขั้นแรกโดยจะแบ่งเป็น

พิกเซลของวัตถุ $G_1 = \{f(x,y) : f(x,y) > T\}$ (2.3)

พิกเซลของพื้นหลัง $G_1 = \{f(x,y) : f(x,y) \le T\}$ (2.4)

โดย F(row,col) คือค่าของพิกเซลที่ตำแหน่งแถวและสดมภ์นั้นๆ

- คำนวณหาค่าเฉลี่ยของพิกเซลแต่ละกลุ่ม

M1 = ค่าเฉลี่ยของ G1

M2 = ค่าเฉลี่ยของ G2

- หาเทรสโฮลด์ใหม่โดยนำมาจากค่าเฉลี่ยของ M1 และ M2

$$T' = (M1 + M2)/2 \tag{2.5}$$

- ย้อนกลับไปทำขั้นตอนที่สองจนกว่าจะได้ค่าเทรสโฮลด์ใหม่ซ้ำเดิม

2.2.5 การสลับค่า(Invert)

เป็นการสลับค่าภายในพิกเซลนั้นๆ โดยขึ้นอยู่กับขอบเขตุในการเก็บค่าของพิกเซลนั้นๆ ว่าเก็บค่ากี่บิท เช่น 8 บิท ขอบเขต คือ 0-255

สมการ $Xinvert = MaxX - Xoriginal$ (2.6)

72664

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพจากการทำเทรสโฮลด์ตัวอักษรที่ได้จะมีสีขาวและพื้นหลังมีสีดำ เพื่อความสะดวกในการวิเคราะห์เราจะทำการสลับค่าของพิกเซลระหว่างตัวอักษรกับพื้นหลังให้ตัวอักษรมีสีดำ (มีค่าพิกเซลเป็น 1) และพื้นหลังมีสีขาว (มีค่าพิกเซลเป็น 0)

(ก) ก่อนการสลับค่า

(ข) หลังการสลับค่า

รูปที่ 2.15 ภาพการทำสลับค่า

2.2.6 การแบ่งออกเป็นส่วน (Segmentation) เป็นวิธีการแยกตัวอักษรออกจากภาพประโยค แบ่งเป็นขั้นตอนได้ดังนี้

การค้นหาแนว (Line Detection)

เป็นการสแกน (Scan) แนวนอนเพื่อแบ่งบรรทัดสำหรับภาพที่ตัดมาจากไฟล์วีดีโอนั้น โดยคำบรรยายข้างใต้ภาพยนตร์ที่ได้อาจมีมากกว่าหนึ่งบรรทัด เริ่มจากการสแกนหาจุดที่เป็นสีดำจุดแรก จากจุดพิกเซลที่อยู่มุมบนซ้ายสุดไปจนสิ้นสุดความยาวของภาพ โดยพิกเซลสีดำจุดแรกที่พบจะถูกบันทึกพิกัดไว้เพื่อเป็นขอบบนของภาพ และจะสแกนต่อไปเรื่อยๆ จนพบแถวแรกที่มีค่าพิกเซลเป็น 0 หรือเป็นสีขาวทั้งแถว แล้วจะทำการบันทึกพิกัดไว้เพื่อเป็นขอบล่างของภาพ ซึ่งจะเท่ากับพิกัดแกน y-1

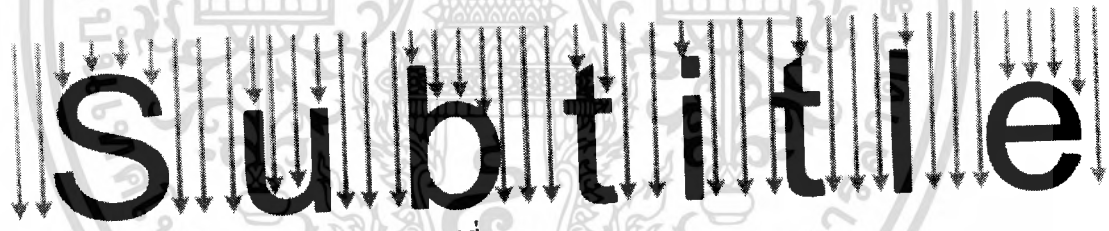
จากนั้นทำการสแกนแนวนอนเพื่อหาพิกเซลจุดที่เป็นสีดำจุดแรกและหาแถวแรกที่เป็นสีขาวทั้งแถวซ้ำ จนถึงมุมล่างขวาสุดของภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 2.16 Line Detection

การค้นหาช่องว่าง (Gap Detection)

เป็นการตรวจกวาดเพื่อแยกตัวอักษรออกจากกัน โดยจะสแกนตามแนวตั้งจากบนสุดจนถึงล่างสุด โดยพิกเซลสีดำจุดแรกที่พบจะเป็นขอบซ้ายสุดของภาพตัวอักษร และจะสแกนต่อไปเรื่อยๆ จนพบคอลัมภ์แรกที่เป็นสีขาวเพื่อเป็นขอบขวาสุดของภาพอักษร ซึ่งจะเท่ากับพิกัดของแกน x-1

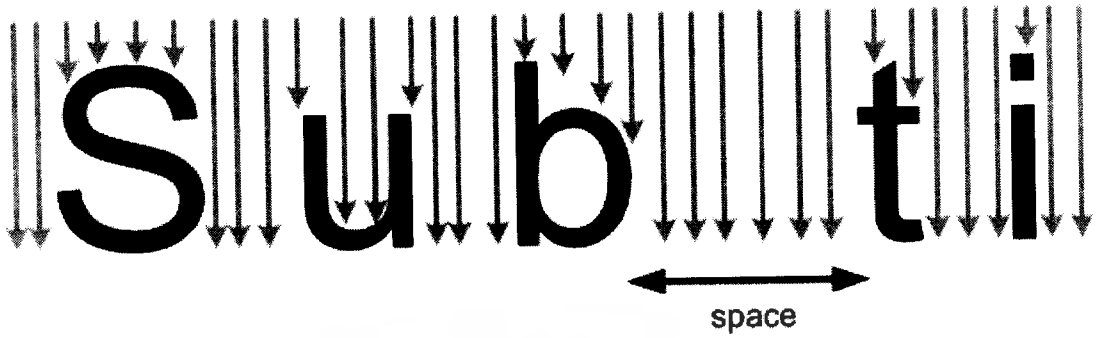


รูปที่ 2.17 Gap Detection

การค้นหาช่องว่างระหว่างตัวอักษร (Character Space Detection)

เป็นการตรวจหาช่องว่างระหว่างคำ โดยพิจารณาว่า ความกว้างจากขอบขวาสุดของอักษรตัวก่อนหน้ากับขอบซ้ายสุดของภาพอักษรถัดมาซึ่งถ้าหากมีความกว้างเกินที่กำหนดไว้ จะถือว่าเกินช่องว่างขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.18 Space Detection

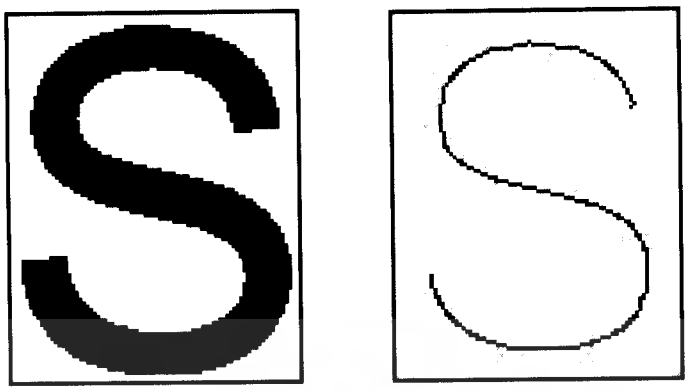
จากนั้นจะทำการตัดตัวอักษรออกจากภาพดั้งเดิมโดยอาศัยค่าที่ได้จากการ Detection ต่างๆ ออกมาทีละตัวเพื่อนำไปทำให้บางและดำเนินการในด้านการจดจำ (Recognize) ต่อไป

รูปที่ 2.19 รูปตัวอักษรที่ทำการแบ่งออกเป็นส่วนจากภาพ

2.2.6 การทำให้บาง (Thinning)

เป็นวิธีการทางการจัดการภาพหนึ่งๆ ที่ช่วยกำจัดสัญญาณรบกวนเพื่อปรับปรุงรูปแบบของตัวอักษรแต่ละตัวให้ดีขึ้น ทำให้สะดวกในการวัดความสูงและความกว้างตัวอักษร และง่ายต่อการนำไปวิเคราะห์ การทำให้บางนั้นใช้วิธีการตรวจกวาดทั้ง 4 ทิศทางซึ่งภาพที่ผ่านการประมวลผลแล้วจะเหลือเฉพาะโครงร่างนั้น โดยใช้หน้าต่าง 3 x 3 เลื่อนไปยังจุดที่ต้องการหาแต่ละจุดแล้วใช้หลักการหาคุณสมบัติของจุดภาพ โดยภาพที่ผ่านกระบวนการนี้แล้วจะเหลือเฉพาะโครงร่าง (หนา 1 พิกเซล)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ก) ก่อนทำ Thinning (ข) หลังทำ Thinning

รูปที่ 2.20 ภาพการทำ Thinning

2.3 การจำแนกตัวอักษร (Character Recognition)

การจำแนกตัวอักษรคือกระบวนการที่ทำให้คอมพิวเตอร์สามารถระบุตัวอักษรหรือข้อความจากรูปภาพได้ โดยการวิเคราะห์ (Analyze) ภาพแล้วให้ความหมาย (Interpret) แล้วทำการตีความหมาย (Translate) ออกมา การจดจำมีกระบวนการหลายขั้นตอน แต่ทางปฏิบัติจะไม่ใช้ทุกกระบวนการเนื่องจากกระบวนการทั้งหมดประมวลผลในเครื่องคอมพิวเตอร์จึงสามารถตัดขั้นตอนบางอย่างออกไปได้

โครงสร้างของระบบโดยทั่วๆ ไปประกอบไปด้วยขั้นตอนการทำงานหลัก 3 ขั้น

รูปที่ 2.21 โครงสร้างทั่วไปของการจำแนกตัวอักษร

ซึ่งในแต่ละขั้นตอนประกอบด้วยรายละเอียดดังต่อไปนี้

1) ขบวนการประมวลผลขั้นต้น (Pre-Processing)

ก่อนที่โปรแกรมจะสามารถบอกได้ว่ารูปภาพที่ส่งเข้าไปประกอบด้วยตัวอักษรอะไรบ้าง ขั้นตอนดังกล่าวนี้มักถูกเรียกรวมกันว่า กระบวนการประมวลผลขั้นต้น (Pre-Processing) ซึ่งเป็นขั้นตอนในการปรับแต่งและจัดเตรียมข้อมูลให้เหมาะสมกับขั้นตอนการรู้จำต่อไป ขั้นตอนเหล่านี้มีความสำคัญต่อประสิทธิภาพโดยรวมของระบบ เพราะหากมีความผิดพลาดเกิดขึ้นในส่วนนี้ ก็จะส่งผลกระทบไปยังส่วนถัดไปของระบบด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนการประมวลผลเบื้องต้นที่สำคัญ ได้แก่

1.1) การกรองข้อมูลแทรกซ้อน (Noise Filtering) การกรองข้อมูลแทรกซ้อนมีจุดประสงค์เพื่อลดทอนส่วนของรูปภาพที่เป็นสิ่งแปลกปลอมอันไม่พึงประสงค์ออกไป โดยข้อมูลแทรกซ้อนที่เกิดขึ้นส่วนใหญ่มักจะมาจากคุณภาพของเอกสารต้นฉบับที่นำมาทำการอ่าน ซึ่งเป็นต้นเหตุสำคัญที่ทำให้ความถูกต้องของโปรแกรมลดลง จึงจำเป็นจัดการกับส่วนเกินเหล่านี้ออกไปให้ได้มากที่สุดเท่าที่จะเป็นไปได้ แต่เรายังไม่มีวิธีการใดที่รับรองได้ว่าสามารถจัดการกับข้อมูลแทรกซ้อนได้ โดยสมบูรณ์ ดังนั้นส่วนการรู้จำก็จะต้องมีความทนทานต่อการแทรกซ้อนเหล่านี้ได้พอสมควร

1.2) การปรับแต่งข้อมูล (Normalization) การปรับแต่งข้อมูลเป็นการปรับภาพตัวอักษรให้อยู่ในรูปแบบที่ระบบต้องการเพื่อนำไปใช้ในขั้น เช่น การปรับขนาดรูปตัวอักษร, การปรับตัวอักษรที่เอียงให้ตรง, การแปลงรูปสีหรือเกรย์สเกลให้เป็นขาวดำ หรือในทางกลับกัน การแปลงรูปขาวดำให้เป็นสีหรือเกรย์สเกล เป็นต้น

1.3) การตัดแบ่งพื้นที่ใช้งาน (Cropping) เป็นการตัดแบ่งพื้นที่เป็นการตัดแยกเอาเฉพาะรูปตัวอักษรออกมาจากภาพ เพื่อส่งให้ขั้นตอนการรู้จำในการระบุว่ารูปตัวอักษรนั้นเป็นรหัสอักษรอะไร หลักการพอสังเขปที่ใช้สำหรับการตัดรูปตัวอักษรโดยทั่วไปจะใช้พื้นที่สีขาว (สีพื้น) รอบรูปเป็นตัวกำหนดขอบเขตในการตัด ซึ่งในขั้นตอนนี้มักจะประสบปัญหาที่ส่งผลกระทบต่ออัตราความถูกต้องของระบบโดยรวมอยู่สองปัญหา ปัญหาแรกคือปัญหาตัวติด เกิดจากรูปของตัวอักษรตั้งแต่สองตัวขึ้นไปมีส่วนที่เชื่อมติดกันทำให้ไม่สามารถแยกตัวอักษรออกจากกันโดยใช้พื้นที่สีขาวรอบๆ ได้ จำเป็นต้องหาอัลกอริธึมพิเศษมาช่วยในการแยกตัวอักษรออกจากกัน ส่วนปัญหาที่สองในทางตรงกันข้าม เป็นปัญหาตัวขาดที่รูปตัวอักษรหนึ่งๆ ถูกแยกออกเป็นส่วนๆ ทำให้เวลาตัดตัวอักษรจากตัวเดียวจะได้เป็นสองตัว ซึ่งก็ต้องหาวิธีการเฉพาะสำหรับมาจัดการอีกเช่นกัน

1.4) การสกัดลักษณะสำคัญ (Feature Extraction) เป็นการดึงเอาลักษณะโครงสร้างพื้นฐานที่สำคัญของตัวอักษรนั้นออกมา โดยโครงสร้างพื้นฐานที่ว่าจะต้องมีการกำหนดไว้ก่อนว่าจะมีอะไรบ้าง มีการนิยามอย่างไร ตัวอย่างเช่น สำหรับภาษาไทยเราอาจกำหนดว่าตัวอักษรภาษาไทยทั้งหมดประกอบด้วยโครงสร้างพื้นฐานคือ เส้นตรง (แนวตั้ง/นอน) เส้นเอียง หัว (วงกลม) ส่วนโค้ง ส่วนเว้า จุดแตกกิ่ง จุดตัด เป็นต้น เมื่อเราสามารถแยกเอาองค์ประกอบของตัวอักษรแต่ละตัวออกมาได้แล้วเราก็นำเสนอรูปภาพของตัวอักษรนั้นในรูปแบบของรายการขององค์ประกอบพื้นฐานต่างๆ แทน ซึ่งจะถูกส่งต่อเป็นอินพุตสำหรับขั้นตอนการรู้จำต่อไป

2) การรู้จำ (Recognition)

ขั้นตอนนี้ถือเป็นหัวใจของระบบ เพราะเป็นส่วนที่จะตัดสินว่ารูปตัวอักษรที่ส่งเข้าไปเป็นรหัสตัวอักษรอะไร เช่นเดียวกับส่วนอื่นๆ ที่มีวิธีการหลากหลายซึ่งนำมาใช้เพื่อให้ได้ผลการทำงาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ที่ดีที่สุด เทคนิคใหม่ๆ ได้ถูกพัฒนาขึ้นมาอย่างต่อเนื่องเพื่อรองรับกับปัญหาที่เกิดจากเอกสารที่มีความสลับซับซ้อนมากขึ้น อย่างไรก็ตามเราพอที่จะจัดแบ่งเทคนิคเหล่านี้ออกเป็นกลุ่มตามแนวทางหลักที่ใช้ในการแก้ปัญหา ถึงแม้บ่อยครั้งที่พบว่ามีความคาบเกี่ยวกันของเทคนิคที่นำมาใช้ระหว่างกลุ่มที่ว่านี้ ทั้งนี้เพราะแต่ละแนวทางก็มีข้อดีข้อเสียที่แตกต่างกัน จึงมีความพยายามที่จะรวมเอาเทคนิคเหล่านี้มาใช้ร่วมกันเพื่อเพิ่มความสามารถของระบบเท่าที่จะเป็นไปได้ ดังนั้นการแบ่งกลุ่มในที่นี้ เป็นการแบ่งที่เน้นความชัดเจนในแง่ของขอบเขตทางทฤษฎีเป็นหลัก

วิธีทางการเข้าคู่รูปแบบ (Template Matching) เป็นวิธีการเข้าคู่รูปแบบเป็นวิธีการแรกๆ ที่มาใช้ในการรู้จำตัวอักษร หลักการโดยทั่วไปคือ จะต้องมีรูปแบบ (template) ที่สร้างขึ้นมาสำหรับอ่านตัวอักษร โดยมีการกำหนดตำแหน่งสำคัญที่สามารถใช้แยกแยะความแตกต่างระหว่างตัวอักษรแต่ละตัว เวลาทำงานก็ให้นำรูปภาพที่ต้องการอ่านไปทาบบนแบบเพื่อวัดความคล้ายคลึงกันของภาพกับตัวแบบ จากนั้นก็ระบุว่าเป็นรหัสตัวอักษรอะไร โดยใช้ค่าผ่านระดับหรือวิธีการบางอย่างในการตัดสิน วิธีการนี้จะค่อนข้างอ่อนไหวต่อข้อมูลแทรกซ้อน ขนาด และการเอียงของตัวอักษร จึงจำเป็นต้องมีขั้นตอนการปรับแต่งข้อมูลที่ดี นอกจากนั้นขั้นตอนการเปรียบเทียบก็ไม่ใช่ว่าสามารถเทียบกันแบบจุดต่อจุดได้ เพราะในทางปฏิบัติตัวอักษรที่ส่งเข้าสามารถมีความแปรปรวนได้หลายรูปแบบ ดังนั้นวิธีการเทียบก็ต้องมีประสิทธิภาพเพียงพอที่จะรองรับกับปัญหาดังกล่าวได้

3) ขบวนการประมวลผลขั้นปลาย (Post-Processing)

หลังจากผ่านขั้นตอนการรู้จำแล้วรูปตัวอักษรที่ถูกส่งเข้าไปจะได้ผลลัพธ์ออกมาเป็นรหัสตัวอักษรซึ่งก็ไม่ได้หมายความว่าเอาท์พุตที่ได้มาจะถูกต้องทั้งหมด ไม่มีผลิตภัณฑ์โอซีอาร์ตัวใด ไม่ว่าจะเป็นภาษาใดก็ตามที่สามารถรับรองความถูกต้องได้ถึง 100 % ดังนั้นเพื่อเพิ่มความถูกต้องให้แก่โปรแกรมจึงได้มีการเสริมส่วนการตรวจสอบและแก้ไขข้อความเข้ามาโดยโปรแกรมส่วนนี้ทำงานเกี่ยวกับการตรวจสอบความถูกต้องของการสะกดคำและไวยากรณ์ภาษา โดยใช้พจนานุกรมมาช่วยในการตรวจสอบคำผิด ซึ่งอาจแก้ไขให้โดยอัตโนมัติหรือแสดงเครื่องหมายบางอย่างเพื่อบอกให้ผู้ใช้ทราบว่าคำดังกล่าวอาจไม่ถูกต้อง ซึ่งผู้ใช้อาจแก้หรือไม่แก้ก็ขึ้นกับการตัดสินใจของผู้ใช้เอง นอกเหนือไปจากการตรวจสอบความถูกต้องระดับคำแล้ว นอกจากนี้บางโปรแกรมความสามารถตรวจสอบไวยากรณ์ในระดับประโยคได้ด้วย

2.3.1 โอซีอาร์ (OCR หรือ Optical Character Recognition)

หมายถึงการรู้จำตัวอักษรโดยการมองเห็นภาพ คือ การอ่านภาพซึ่งในที่นี้คือตัวอักษร โดยนำมาเปรียบเทียบกันต้นแบบที่ได้จดจำไว้ก่อนแล้วประมวลผลว่ารูปแบบตัวอักษรนั้นเป็นตัวอักษรอะไร โอซีอาร์ทำได้หลายวิธี เช่น วิธีวิเคราะห์ทางโครงสร้าง (Structure Analysis) , วิธีทางโครงข่าย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประสาทเทียม (Neural Network) , วิธีวิเคราะห์ทางสถิติ (Statistic Approach) ,วิธีการเข้าคู่รูปแบบ (Template Matching) ฯลฯ

2.3.2 การรู้จำรูปแบบ (Pattern Recognition)

โปรแกรมประเภทโอซีอาร์เป็นโปรแกรมที่ประยุกต์ความรู้ด้านการรู้จำรูปแบบ (Pattern Recognition) ซึ่งต้องอาศัยเทคนิคการวิเคราะห์แยกแยะและการระบุชนิด โดยอินพุตของโปรแกรมมักจะต้องผ่านกระบวนการย่อยๆ อีกหลายขั้นขึ้นอยู่กับการประยุกต์ใช้ ได้แก่ การกรองข้อมูลแทรกซ้อน การปรับแต่งข้อมูล การเลือกตัดแบ่งพื้นที่ใช้งาน การสกัดลักษณะสำคัญ การจำแนกแยกแยะ การรู้จำ และอื่นๆ โดยในแต่ละขั้นตอนก็สามารถจะมีวิธีการได้อีกหลากหลายวิธี ขึ้นอยู่กับตัวข้อมูลและการประยุกต์ใช้

นอกจากโอซีอาร์แล้ว งานประยุกต์ที่แตกแขนงออกมาจากการรู้จำแบบได้แก่ การรับรู้ภาพของคอมพิวเตอร์, การวิเคราะห์การเกิดแผ่นดินไหว, การแยกแยะและวิเคราะห์สัญญาณเรดาร์, การรู้จำใบหน้า, การรู้จำเสียงพูด, การเข้าใจเสียงพูด, การระบุลายพิมพ์นิ้วมือ, การวินิจฉัยทางการแพทย์ และอื่นๆ

2.3.3 วิธีการวิเคราะห์ทางโครงสร้าง (Structural Analysis)

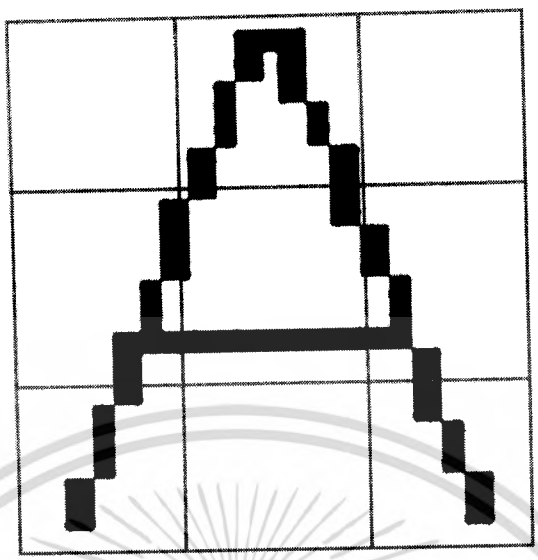
วิธีการวิเคราะห์ทางโครงสร้าง คือ การวิเคราะห์โครงสร้างตัวอักษร โดยถือว่าตัวอักษรทุกตัวประกอบด้วยองค์ประกอบพื้นฐาน ซึ่งได้มาจากการสกัดลักษณะสำคัญ โดยรูปภาพอินพุตที่ได้มาจากขั้นตอนการสกัดลักษณะ คือผ่านการ ไบนาไรซ์ (binarize) และการทำให้บาง (thinning) แล้ว จะถูกส่งเข้าไปในส่วนการรู้จำเฉพาะของแต่ละตัวอักษร ซึ่งจะตัดสินโดยการดูที่รูปแบบการเชื่อมต่อขององค์ประกอบต่างๆ เข้าเป็นตัวอักษรนั้น ซึ่งได้ผลลัพธ์จะได้ผลลัพธ์ออกมาเป็นค่าของแต่ละตัวอักษร เมื่ออินพุตได้ผ่านส่วนการรู้จำครบทุกตัวแล้ว ก็นำเอาผลลัพธ์ที่ได้ทั้งหมดมาเป็นฐานข้อมูล เพื่อใช้เปรียบเทียบกันว่าตัวอักษรที่นำมาพิจารณาเป็นตัวอักษรใด

การวิเคราะห์ทางโครงสร้างโดยใช้หลักของโทโพจี (Topology Recognize)

2.3.3.1 การแบ่งภาพอักษรออกเป็น 9 ส่วน

โดยจะนำแบ่งภาพออกเป็น 9 ส่วน แต่ละส่วนจะถูกนำมาพิจารณาเพื่อหาคุณสมบัติทางโทโพโลจีของจุดภาพแล้วนำมากำหนดเป็นรหัสขนาด 4 บิต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.22 รูปตัวอักษร "A" ที่นำมาคำหาโทโพโลจี

2.3.3.2 การวิเคราะห์โดยใช้คุณสมบัติทางโทโพโลจีของภาพ

จะใช้วิธีแบ่งส่วนของตัวอักษรออกเป็น 9 ส่วน เพื่อพิจารณาคุณสมบัติทางโทโปโลยีของจุดภาพในแต่ละส่วน โดยจะพิจารณาคุณสมบัติทางโทโปโลยีของจุดภาพเพียง 4 แบบ คือ จุดปลาย (End พิกเซล), จุดต่อ (Connect พิกเซล), จุดแยก (Branch พิกเซล) และจุดตัด (Cross พิกเซล)

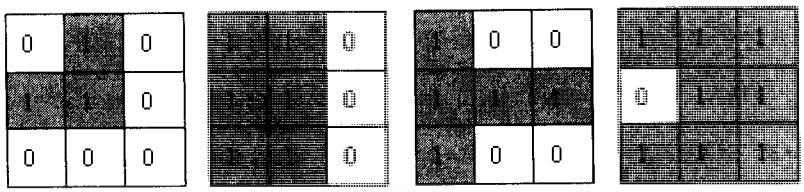
N8	N1	N2
N7	1	N3
N6	N5	N4

รูปที่ 2.23 ตัวอย่างของจุดภาพที่จะนำมาใช้วิเคราะห์

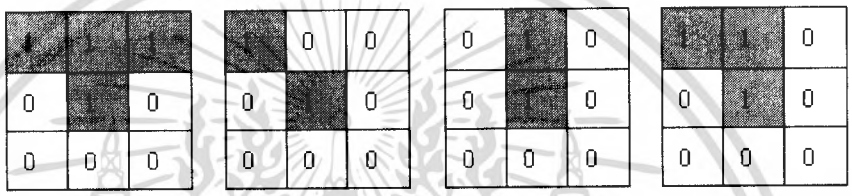
โดยจะพิจารณาเฉพาะค่า พิกเซล ของภาพอักษรที่เป็นสีดำซึ่งมีค่า พิกเซล ณ ตำแหน่งนั้นเป็น 1 เป็นพิกัดตรงกลาง แล้วนับจำนวนการเปลี่ยนแปลงของค่า พิกเซล ข้างเคียง หากเกิดการเปลี่ยนแปลงค่าจาก 0 → 1 หรือจาก 1 → 0 โดยจะพิจารณาจาก N1 → N2 → N3 → N4→ N5 → N6 → N7 → N8→ N1 ซึ่งลักษณะทางโทโพโลจีของจุดภาพมีลักษณะดังต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

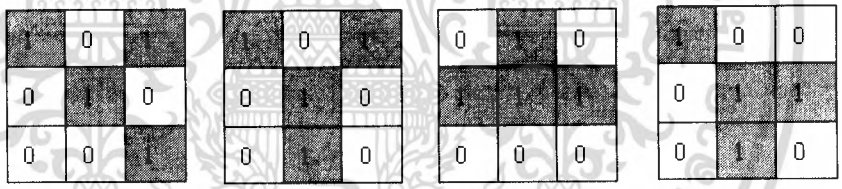
จุดต่อ (โทโพโลจี = 4)



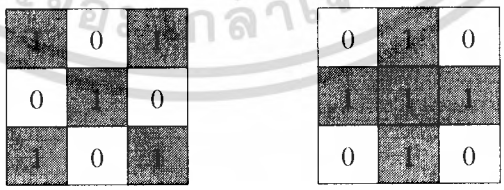
จุดปลาย (โทโพโลจี = 2)



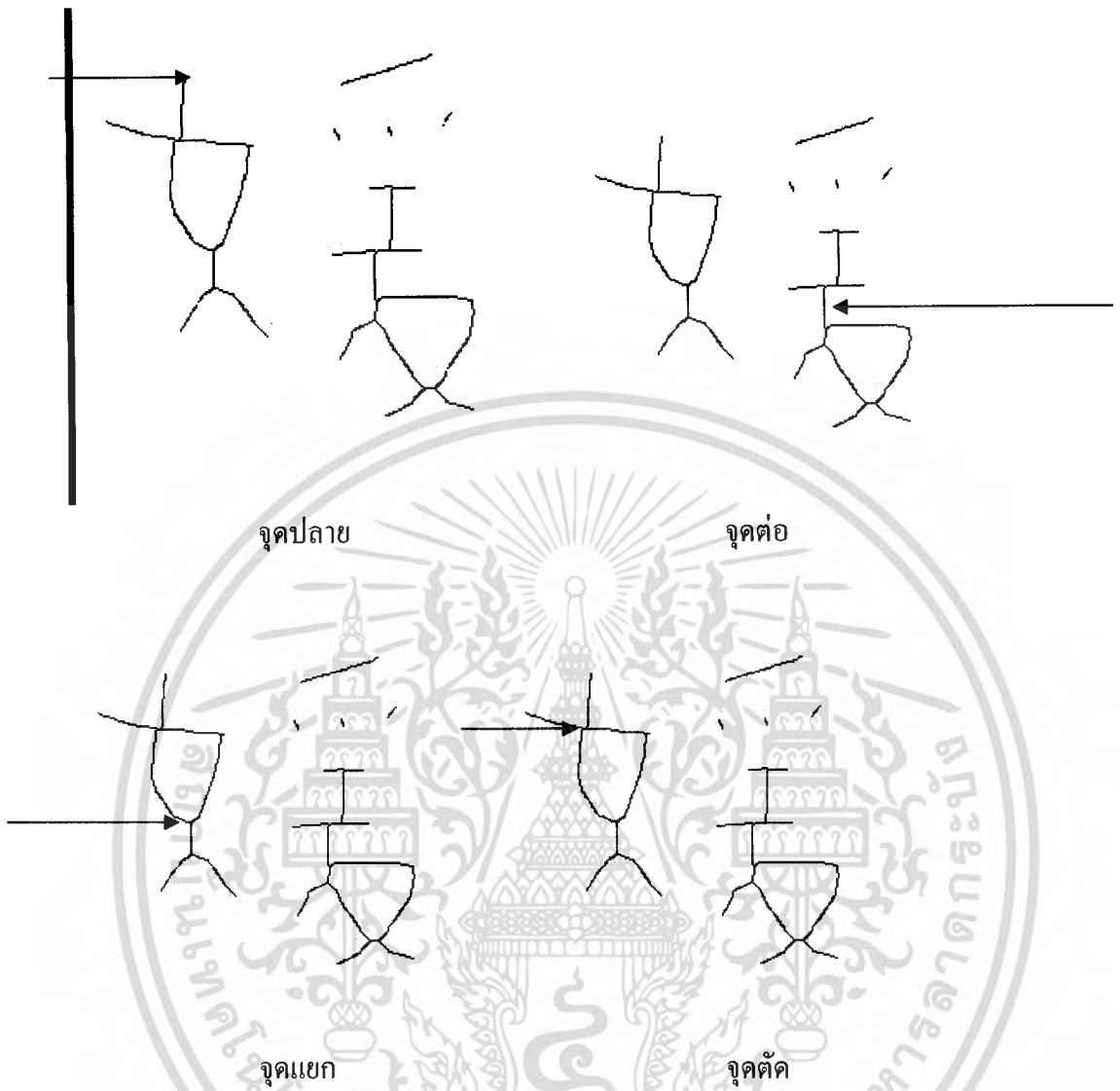
จุดแยก (โทโพโลจี = 6)



จุดตัด (โทโพโลจี = 8)



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

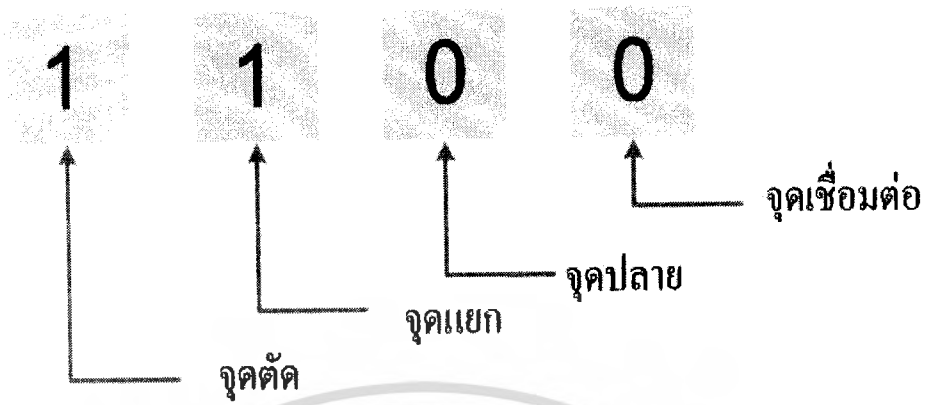


รูปที่ 2.24 ตัวอย่างลักษณะของคุณสมบัติโทโปโลยีในภาพอักษร

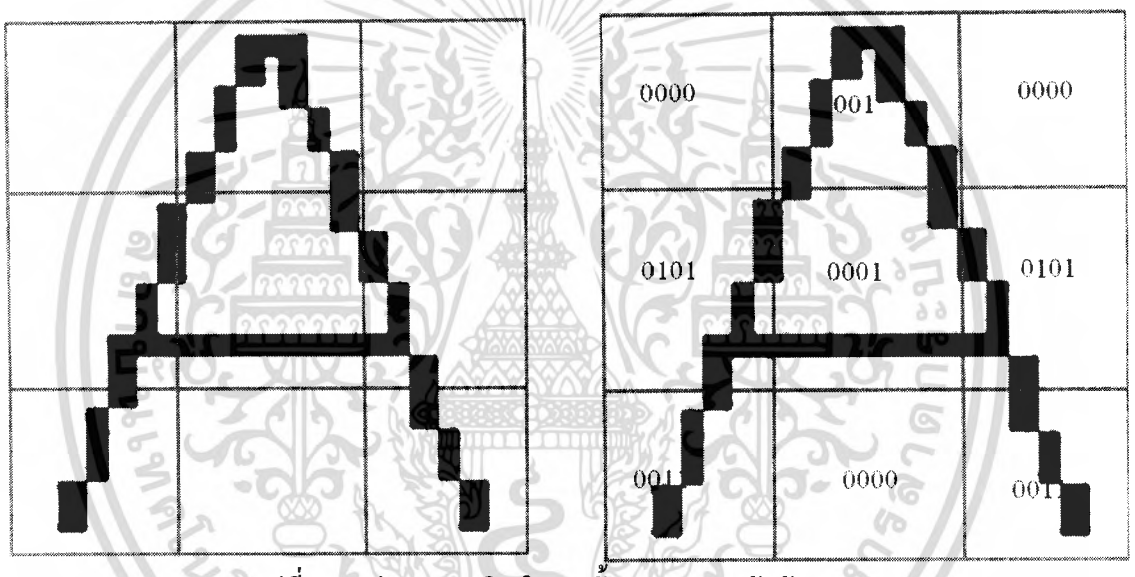
2.3.3.3 การวิเคราะห์หาค่ารหัส 4 บิตลงในภาพอักษร

เพื่อหาว่าในแต่ละส่วนของภาพอักษรที่ถูกแบ่งออกเป็น 9 ช่องนั้น มีคุณสมบัติของแยก,จุดปลาย,จุดตัดหรือ จุดเชื่อมต่ออยู่หรือไม่ ถ้าส่วนนั้นมีคุณสมบัติใดก็จะให้บิตที่แทนคุณสมบัตินั้นมีค่าเป็น 1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.25 ค่ารหัส 4 บิต จุดแยก, จุดปลาย, จุดตัด, จุดต่อ



รูปที่ 2.26 ค่ารหัส 4 บิต ในจุดทั้ง 9 ของภาพตัวอักษร "A"

2.3.3.4 การจัดเก็บลงฐานข้อมูลของการรู้จำ

ลักษณะการจัดเก็บลงฐานข้อมูลจะจัดเก็บ 9 บิตต่อ 1 ภาพอักษร โดยจะนำรหัส 4 บิตของแต่ละส่วนของภาพอักษรคูณด้วยค่าประจำตำแหน่งของเลขฐาน 2 คือ

บิตที่ 0 คูณด้วย 2^0

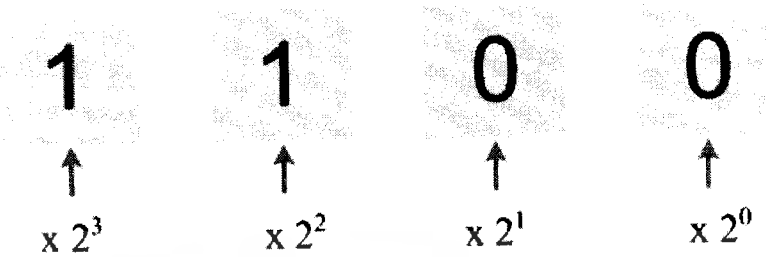
บิตที่ 1 คูณด้วย 2^1

บิตที่ 2 คูณด้วย 2^2

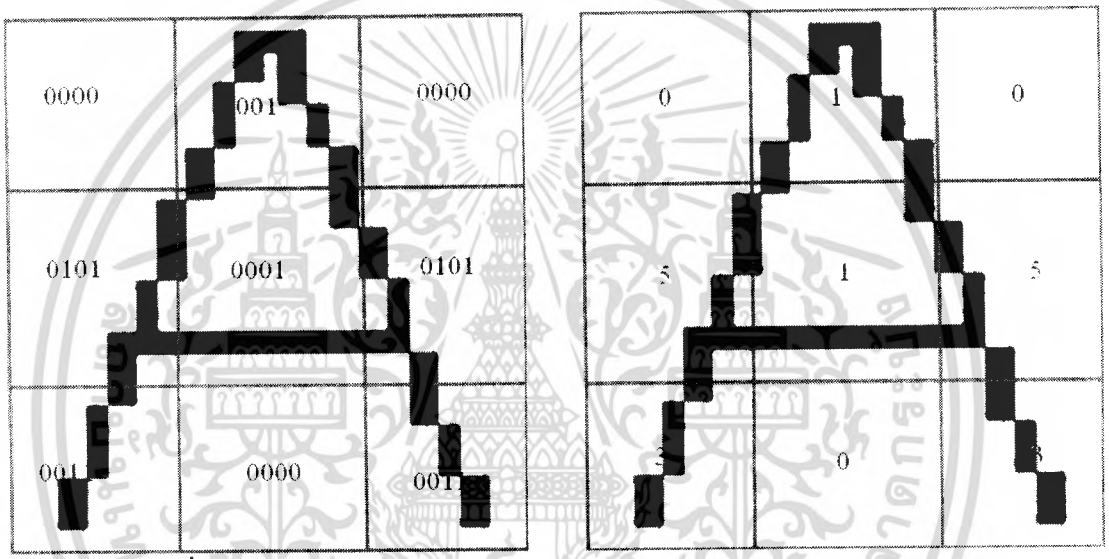
บิตที่ 3 คูณด้วย 2^3

แล้วนำค่าที่ได้จากแต่ละส่วนเก็บลงในฐานข้อมูล ซึ่งจะเก็บเป็นอาร์เรย์เพื่อการเปรียบเทียบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.27 ค่าประจำหลักของแต่ละจุดในค่ารหัส 4 บิต



รูปที่ 2.28 ค่ารหัส 4 บิตลงในภาพอักษรหลังจากคูณด้วยค่าประจำตำแหน่ง

2.4 เครื่องมือที่ใช้พัฒนาและไลบรารี่ที่เกี่ยวข้อง

2.4.1 การประมวลผลภาพขั้นสูงของจาวา (Java Advance Imaging)

เป็น เอพีไอ (API หรือชื่อเต็มคือ Application Programmer's Interfaces) ในการจัดการรูปภาพในระดับธรรมดาและระดับสูง เพื่อความสะดวกในการพัฒนาด้วย แอพพลิเคชั่น และ แอปเพล็ต ซึ่งรวบรวมฟังก์ชั่นที่จำเป็นสำหรับใช้งาน เกี่ยวกับรูปภาพ เช่น การเข้าถึงในระดับพิกเซล, การโหลดรูปภาพมาแสดงในตัวโปรแกรม, แก้ไขปรับแต่งรูปภาพ ซึ่งเหนือกว่าคลาสของจาวาสำหรับภาพ2มิติ ซึ่งถูกรวมเข้าไว้ในจาวาเอพเพท (Java applet) และ แอพพลิเคชั่น (application)

เจเอไอ เอพีไอ ก้าวหน้าไปจากฟังก์ชั่นและ จาวา ธรรมดา โดย เอพีไอ ตัวนี้ได้รวมความสามารถอาทิ เช่น image tilling, regions of interest, deferred execution และชุดคำสั่งของการประมวลผลภาพ เช่น การปรับภาพแบบกระทำต่อจุด, แบบกระทำเป็นพื้นที่ และชุดคำสั่งการทำ frequency domain ในงานวิจัยนี้จะ class PlanarImage เป็นหลักในการจัดการภาพจากไฟล์ เอวีไอ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เจเอไอ เอพีไอ มีข้อดีสำหรับผู้ชำนาญด้านประมวลผลภาพ ดังนี้

เจเอไอ เอพีไอ ขยายขอบเขตของ จาวา Platform โดยให้มีฟังก์ชั่นการทำงานที่ซับซ้อนด้านการประมวลผลภาพซึ่งถูกรวมเข้าไว้ในจาวาแอปเพล็ตและแอพพลิเคชั่น

เจเอไอ เอพีไอ ก้าวหน้าไปจากฟังก์ชั่นเดิมโดยมีประสิทธิภาพสูง ใช้งานผ่านแพลตฟอร์มอย่างอิสระและเพิ่มโครงร่าง (Framework) การประมวลผลภาพเข้าไป

เนื่องด้านโครงสร้างโปรแกรมที่ง่ายแก่การใช้งาน ทำให้งานหนักๆ ที่ต้องการสร้างซอฟแวร์ง่ายลง ทำให้เวลาที่ใช้ในการพัฒนาโปรแกรมลดลง

เนื่องจากเอพีไอตัวนี้สร้างมาบนศูนย์กลางเครือข่ายของจาวาแพลทฟรอม ทำให้นักพัฒนาสามารถใช้มันโปรแกรมร่วมกันเพื่อการทำการประมวลผลภาพระดับสูงหรือการทำวิเคราะห์ผ่านเครือข่าย

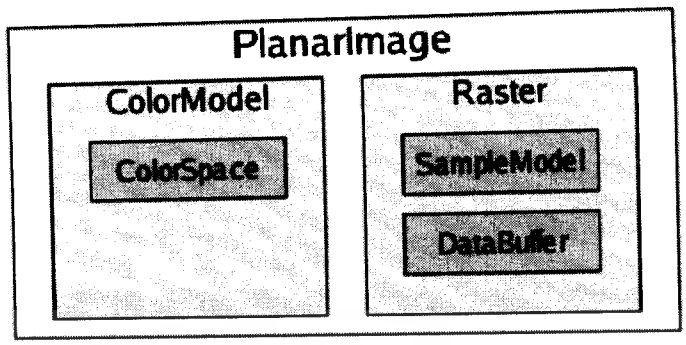
2.4.1.1 ขั้นตอนการดำเนินงานและตัวดำเนินการ เจเอไอ

เจเอไอ เอพีไอ สามารถนำมาใช้สำหรับการทำกระบวนการประมวลผลรูปภาพ (image processing) และการแสดงผล แม้ว่าจะไม่มีความสามารถของซอฟต์แวร์ประมวลผลรูปภาพทั้งหมด แต่ฟังก์ชั่นและการที่สามารถขยายการใช้งาน ทำให้เจเอไอเป็นทางเลือกที่ดีทางหนึ่งในการพัฒนาชุดขั้นตอนคำสั่งการประมวลผลรูปภาพ

ชุดขั้นตอนคำสั่งการประมวลผลรูปภาพโดยทั่วไปจะต้องการความสามารถในการจัดการข้อมูลของพิกเซลโดยภาพในเจเอไอ อาจจะเป็นได้ทั้งแบบ หลายมิติ และ อาจจะมีพิกเซลที่มีค่าเป็นจำนวนเต็มหรือจำนวนทศนิยม พิกเซลอาจจะรวมไว้ด้วยวิธีการต่างๆ กันหรือกระจายออกมาอยู่ใน อาเรย์ (array) และรูปแบบของสีนั้นก็สามารถใช้ได้หลายรูปแบบเช่นกัน

ในการจะแสดงข้อมูลภาพที่หลากหลายก็จะมีคลาสต่างๆกันออกไปมาจัดการดังนี้ พลานาอิมเมจ (PlanarImage) : คลาสพื้นฐานของการแสดงภาพในเจเอไอ สามารถแสดงภาพโดยมีความยืดหยุ่นมากกว่าคลาส บัฟเฟออิมเมจ (BuferedImage) ของจาวา โดยทั้งบัฟเฟออิมเมจและพลานาอิมเมจเองนั้นก็จะมีคลาสที่ต่างๆ กันออกไปเพื่อให้เกิดความยืดหยุ่นในการแสดงภาพโดยมี แรสเตอร์ (Raster) เป็นอินสแตนส์ (instance) ที่ใช้ในการเก็บค่าของพิกเซลตามรูปแบบที่กำหนดโดยอินสแตนส์ของแซมเปิ้ลโมเดล (SampleModel) ของพลานาอิมเมจ นอกจากนี้ยังมีคัลเลอร์โมเดล (ColorModel) มาเกี่ยวข้องด้วย ซึ่งบรรจุด้วยอินสแตนส์ของคัลเลอร์สเปซ (ColorSpace) ใช้ในการระบุว่าค่าในพิกเซลจะแปลงเป็นค่าสีอย่างไร ดังรูปที่ 1 แสดงถึงคลาสที่ใช้ในการประกอบกันเป็นอินสแตนส์ของ พลานาอิมเมจ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.29 โครงสร้างของพลานาอิมเมจ

พลานาอิมเมจนั้นจะอ่านได้อย่างเดียว โดยจะสร้างขึ้นมาได้และสามารถอ่านค่าของแต่ละพิกเซลออกมาได้หลายวิธี แต่จะไม่มีการอนุญาตให้แก้ไขค่าของพิกเซล พลานาอิมเมจอาจมีจุดพิกัดเริ่มต้นนอกเหนือจาก พิกัด (0,0) ได้ หรือแม้กระทั่งค่าที่เป็นค่าติดลบ

ไทล์อิมเมจ (TiledImage) : เป็นคลาสย่อยของพลานาอิมเมจซึ่งสามารถนำมาอ่านและเขียนข้อมูลภาพได้

เรนเดอร์ (RenderedOp) : คลาสย่อยของพลานาอิมเมจ อีกตัวหนึ่ง ซึ่งแสดงเป็นในลักษณะของ การสร้างภาพเป็นลูกโซ่ (redered imaging chain) ซึ่งอนุญาตให้มีการประมวลผลภาพเป็นลำดับตามที่กำหนดและนำไปใช้กับภาพเดียวหรือหลายๆ ภาพได้

ในการสร้างภาพขึ้นมาด้วยเจเอไอ จะเป็นไปตามขั้นตอนดังนี้

1. สร้างข้อมูลของภาพเก็บไว้ในอาเรย์ (array) บนหน่วยความจำอาเรย์นี้ต้องเป็นชนิดมิติเดียว ซึ่งหากสร้างมาเป็นแบบหลายมิติก็ยังสามารถแปลงในภายหลังได้

2. สร้างอินสแตนส์ที่เป็นคลาสย่อยชนิดแบบคอนครีต (concrete subclass) ของ DataBuffer โดยใช้คอนสตรักเตอร์ (constructor) ของตัวมัน และ อาเรย์ข้อมูลภาพ

3. สร้างอินสแตนส์ของแซมเปิ้ลโมเดล ที่มีชนิดข้อมูลเหมือนกับของ บัฟเฟอร์ข้อมูล (DataBuffer) และกำหนดขนาดที่ต้องการ โดยในการกำหนดอาจจะใช้เมทธอด (method) จาก คลาส แรสเตอร์แฟคทอรี (RasterFactory)

4. สร้างอินสแตนส์ของคัลเลอร์โมเดล ที่เข้ากันได้กับแซมเปิ้ลโมเดล ที่ใช้อยู่ โดยใช้เมทธอดแบบคงที่ (static method) PlanarImage.createColorModel ที่มีอาร์กิวเมนต์ (argument) เป็นโมเดลตัวอย่าง

5. สร้างอินสแตนส์ของ ไรต์เอเบิลแรสเตอร์ (WritableRaster) โดยใช้แซมเปิ้ลโมเดลและอาร์เรย์ของข้อมูลภาพ ใช้เมธอด RasterFactory.createWritableRaster

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6. สร้างอินสแตนส์ของไทล์อิมเมจ ซึ่งเป็นภาพที่อนุญาตให้เขียนได้ โดยใช้ แซมเปิ้ลโมเดล, คัลเลอร์โมเดล และ ขนาดมิติ

7. สร้างความสัมพันธ์อินสแตนส์ของแรสเตอร์กับภาพโดยใช้เมทธอด setData จากคลาสไทล์อิมเมจ

8. ทำการประมวลผล แสดงผล หรือบันทึกภาพ โดยกระทำกับอินสแตนส์ของไทล์อิมเมจ

2.4.1.2 ตัวดำเนินการของเจเอไอ

เจเอไอ เอพีไอ มีอยู่หลายตัวดำเนินการที่สามารถนำมาใช้เพื่อให้เขียนโปรแกรมน้อยลง ตัวดำเนินการเหล่านั้นจำเป็นไปตามแนวคิดของการสร้างภาพเป็นลูกโซ่ ซึ่งแต่ละขั้นตอนจะสามารถกำหนดได้แต่จะดำเนินการจริงๆเมื่อต้องการเท่านั้น

ตัวดำเนินการเหล่านั้นสามารถสร้างขึ้นมาได้ง่ายๆ คือ เริ่มจากสร้างอินสแตนส์ของพารามิเตอร์บล็อก (ParameterBlock) ซึ่งพื้นฐานเป็นเวคเตอร์ (vector) ของข้อมูลที่จะถูกนำมาใช้สำหรับการดำเนินการ จากนั้นเมธอดคงที่ create ของคลาสเจเอไอจะเริ่มการดำเนินการ เมธอดนี้จะถูกเรียกมาเป็นชื่ออาร์กิวเมนต์ของการดำเนินการนั้นและอินสแตนส์ของพารามิเตอร์บล็อกและส่งกลับมาเป็นอินสแตนส์ของเรนเดอร์ออพ (RenderedOp) ซึ่งสามารถนำมาจัดการต่อให้กลายเป็นพลานาอิมเมจได้ หรืออาจจะใส่รูปดั้งเดิมลงไปในอินสแตนส์ของพารามิเตอร์บล็อกเป็นพารามิเตอร์หนึ่งลงไปในเมธอด addSource ส่วนพารามิเตอร์อื่นๆ สามารถเพิ่มลงไปได้ด้วยเมธอด แอด (add)

ตัวดำเนินการสเกล (Scale Operator)

ตัวนำเนินการสเกลจะทำการสเกลรูปภาพโดยให้ผลออกมาเป็นรูปภาพที่สเกลแล้วโดยอาจจะนำมาใช้ในการเคลื่อนย้ายรูปภาพก็ได้ ในการใช้ตัวดำเนินการนี้จะต้องการพารามิเตอร์บล็อก และเพิ่มรูปต้นฉบับ เลขค่าทศนิยมสองค่าไปยังสเกล เอ็กซ์ (X) และ วาย (Y) และอีกสองค่าทศนิยมที่สัมพันธ์กับการเคลื่อนย้ายในเอ็กซ์และวายของพิกเซลของรูปภาพ เมื่อทำการสเกลภาพพิกเซลต้องถูกแทรกเพิ่มเข้าไป ดังนั้นจึงต้องเพิ่มอินสแตนส์ของคอนครีตซับคลาสของ

javax.media.jai.Interpolation เข้าไปด้วย ตัวอย่างการใช้งาน

```
float scale=2.0f;
ParameterBlockpb= new ParameterBlock();
pb.addSource(image);
pb.add(scale);
pb.add(scale);
pb.add(0.0f);
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
pb.add(0.0f);
pb.add(new InterpolationNearest());
PlanarImagescaledImage=JAI.create("scale", pb);
```

ตัวดำเนินการการตัดส่วน

การดำเนินการจะเป็นการเลือกส่วนจากต้นฉบับโดยต้องการพิกัดของจุดเริ่มต้นและขนาดความกว้าง และความสูง

```
ParameterBlock pb2 = new ParameterBlock();
pb2.addSource(jaiImage);
pb2.add(100.0f);
pb2.add(100.0f y);
pb2.add(200.0f);
pb2.add(135.0f);
output = JAI.create("crop",pb2,null);
```

ตัวดำเนินการการเคลื่อนย้าย (Translate Operator)

การดำเนินการการเคลื่อนย้ายจะทำการเคลื่อนย้ายภาพต้นฉบับไปยังจุดที่ต้องการ โดยตัวนำเนินการนี้ต้องการค่าทศนิยมสองค่า ที่บอกถึงพิกัดเอ็กซ์และวาย ที่จะย้ายไป มีตัวอย่างที่จะทำให้ภาพเลื่อนไปด้านเอ็กซ์และวายด้านละ 10 พิกเซล ดังนี้

```
pb = new ParameterBlock();
pb.addSource(overlayS);
pb.add((float) 10.0f);
pb.add((float) 10.0f);
PlanarImage overlay = JAI.create("translate", pb, null);
```

ตัวดำเนินการการรวมแบนด์

ตัวดำเนินการการรวมแบนด์จะใช้ในการรวมแบนด์จากภาพที่รับเข้าไปออกมาได้เป็นผลลัพธ์คือภาพที่จะมีแบนด์เดียว โดยดังตัวอย่างจะเป็นการนำภาพที่รับเข้ามาสร้างเป็นภาพใหม่โดยที่ภาพใหม่จะมีแค่แบนด์เดียวซึ่งเป็นค่าเฉลี่ยของทั้ง3แบนด์

```
double[][] matrix = {{ 0, 0, 0, 0 }};
ParameterBlock pb3 = new ParameterBlock();
pb3.addSource(output2);
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
pb3.add(matrix);
output3 = JAI.create("bandcombine", pb3, null);
```

2.4.1.3 การเข้าถึงข้อมูลของภาพ (Image data access)

บ่อยครั้งที่เราต้องการเข้าถึงค่าในแต่ละพิกเซลของภาพเพื่อที่จะดำเนินการบางอย่างกับรูปนั้น ซึ่งจำเป็นต้องการเข้าถึงทุกๆพิกเซลของภาพ วิธีที่ง่ายที่สุดวิธีหนึ่งที่จะสามารถใช้ในการเข้าถึงทุกพิกเซลคือการใช้ อิเทอเรเตอร์ (iterators) อิเทอเรเตอร์จะยินยอมให้เข้าถึงพิกเซลตามตำแหน่งที่ระบุได้ ดังเช่น อินสแตนส์ของ เรกท์อิเทอ (RectIter) จะไล่ไปตามแต่ละแถว จากแถวบนสุดไปจนถึงแถวล่างสุด โดยจะกระเถิบไปยังพิกเซลถัดไปโดยอัตโนมัติ ทำให้สามารถอ่านทุกๆพิกเซลในแบนด์นั้น อิเทอเรเตอร์อีกชนิดหนึ่งคือ แรนดอมอิเทอเรเตอร์ (RandomIter) จะยอมให้เข้าถึงแต่ละพิกเซลโดยตรงโดยใช้การกำหนดพิกัดเอ็กซ์และวาย

อิเทอเรเตอร์ทั้งสองแบบจริงๆแล้วคืออินเตอร์เฟส (interfaces) ซึ่งสามารถจะสร้างมาได้ด้วยเมธอดในคลาส RectIterFactory และ RandomIterFactory ตามลำดับ เมธอด แฟคตอรี (Factory) ต้องการสองอาร์กิวเมนต์ หนึ่งคือ พลานาอิมเมจ และ หนึ่งอินสแตนส์ของ เรคแทงเกิ้ล (Rectangle) ซึ่งเอามาใช้ระบุส่วนย่อยของภาพที่ต้องการจะนำไปประมวลผล ถ้า ไม่มีค่า แทนอินสแตนส์ของเรคแทงเกิ้ล ภาพทั้งภาพจะถูกนำมาประมวลผล

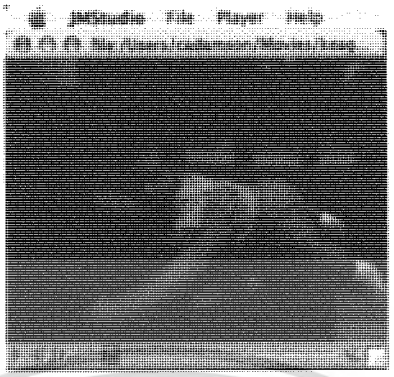
2.4.2 การจัดการด้านมัลติมีเดียของจาวา (JMF หรือ Java Media Framework) เป็นเอพีไอ ของจาวาในการจัดการทางเวลาสามารถทำการการสั่งเก็บภาพ, ย้อนกลับ, จัดความต่อเนื่องและการแปลงโค้ดของ มีเดียได้หลายรูปแบบได้ ใช้สำหรับข้อมูลเสียง, วิดีโอ และสื่อประเภทที่เกี่ยวเนื่องกับเวลาซึ่งเจเอมเอฟถูกเพิ่มเข้าไปในจาวาแอปเพล็ตและแอพพลิเคชั่นอยู่แล้วเช่นเดียวกับเจเอไอซึ่งจะใช้ในการสร้างตัวเพลเออ

เจเอมเอฟรุ่นแรกถูกพัฒนาขึ้นโดยความร่วมมือกันระหว่างบริษัทซันไมโครซิสเตมส์ (Sun Microsystems), บริษัทซิลิคอนกราฟฟิก (Silicon Graphics) และบริษัทอินเทล (Intel) และได้ออกเป็น เจเอมเอฟ 1.0 ในปี 1997 ส่วนเจเอมเอฟ 2.0 ถูกพัฒนาขึ้นโดยบริษัท Sun และ IBM ออกมาในปี 1999 โดยเพิ่มฟังก์ชั่นการสั่งเก็บภาพ, การจัดการรับส่งข้อมูล, การรวมโคเดค, และการแปลงโคด เข้าไป

เจเอมเอฟ 2.0 เดิมจะมีเอมพี 3 (MP 3) การถอดเข้ารหัส และ การเข้ารหัส ด้วยซึ่งถูกเอาออกในปี 2002 และทำการเพิ่ม พลักอิน (plug-in) ของ เอมพี 3 เพลแบคโอนลี่ (playback-only) เข้าไปในปี 2004

ไบนารีของเจเอมเอฟเป็นลิขสิทธิ์ของลูกค้าและส่วนซอสโคด (source code) ต่างๆ เป็นลิขสิทธิ์ของลิขสิทธ์ของกลุ่มซัน (SCSL หรือ Sun Community Source Licensing)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.30 ไฟล์ภาพยนตร์โดยเจเอมสตูดีโอ (JMStudio) ประยุกต์จากเจเอมเอฟ

แนวความคิดในการออกแบบ

เจเอมเอฟ ใช้เปลี่ยนตัวสื่อให้กลายเป็นดาต้าซอส (Datasource) ใช้เรียกสื่อที่ถูกอ่านเข้าไปเข้าไปในเจเอมเอฟ และดาต้าซิ้งค์ (DataSinks) ใช้เรียกสื่อที่ถูกส่งออกไป โดยเจเอมเอฟไม่ได้ถูกออกแบบให้นักพัฒนาสามารถเข้าถึงรูปแบบข้อมูลเพื่อแก้ไขข้อมูล ในทางกลับกันเราให้สื่อถูกแสดงในรูปของ ซอส (source) เช่น รูปแบบตำแหน่ง (.url) ซึ่งสามารถสั่ง อ่าน, เล่น, ประมวลผล หรือ ส่งออก แต่ก็ไม่ได้หมายความว่าทุกโคเดคจะสามารถนำมาประมวลผลและแปลงโคดได้

2.4.2.1 การควบคุมตำแหน่งเฟรม (FramePositioningControl)

การควบคุมตำแหน่งเฟรม() เป็นเมธอดในการจัดการตำแหน่งของเฟรมวีดีโอสำหรับตัวเพลเออ

เฟรมวีดีโอแต่ละเฟรมจะมีหมายเลขเฟรมของตัวเองที่ไม่ซ้ำกัน โดยที่เฟรมหลังจะมีหมายเลขเฟรมมากกว่าเฟรมก่อนหน้า เฟรมแรกจะมีหมายเลขเฟรมเป็น 0 และเวลาเริ่มที่ 0 กระบวนการ ทำให้คล้องจองกัน หมายเลขเฟรมกับเวลา โดยใช้ เมธอดแปลงเฟรมไปเป็นเวลา mapFrameToTime() และ แปลงเวลาไปเป็นเฟรม mapTimeToFrame()

ในการซีค (seek) เฟรมและการข้ามไปยังเฟรมใหม่ เวลาของตัวเพลเออจะเปลี่ยนไปตามเวลาของเฟรมนั้นด้วย เมธอดการเซตเวลา

Time	mapFrameToTime(int frameNumber) เปลี่ยนหมายเลขเฟรมเป็นเวลา
int	mapTimeToFrame(Time mediaTime)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

	เปลี่ยนเวลาเป็นหมายเลขเฟรม
int	seek(int frameNumber) ไปยังเฟรมที่มีหมายเลขเฟรมที่ต้องการ
int	skip(int framesToSkip) จากเฟรมปัจจุบันข้ามไปยังหมายเลขเฟรมที่ต้องการ

ซีค (seek) เมื่อตัวเพลเออต้องการจะข้ามไปยังเฟรมหมายเลขใดๆ เฟรมนั้นจะถูกส่งค่าคืนกลับมาโดยเวลาของตัวเพลเออจะถูกเปลี่ยนค่าไปตามเฟรมที่ส่งกลับมาใหม่นั้นด้วยเมธอด MediaTimeSetEvent() จะถูกเรียกใช้กับการเริ่มหรือการหยุดเล่นของตัวเพลเออ เมธอดเซทเวลามีเดีย setMediaTime() จะเปลี่ยนสถานะของตัวเพลเออที่อยู่ในสถานะเริ่มต้นและเมธอดตรวจสอบความเปลี่ยนแปลง TransitionEvents() จะทำให้ตัวเพลเออทำงาน

ค่าตัวแปร : จำนวนของเฟรม (frameNumber) เป็นหมายเลขเฟรมที่จะข้ามไป

ค่าที่ส่งกลับ : เฟรมที่ต้องการข้ามไป

สคิบ (skip) เป็นการกระโดดข้ามจากเฟรมปัจจุบันไปยังเฟรมที่ต้องการ เมื่อตัวเพลเออที่ต้องการข้ามการเล่นจากเฟรมปัจจุบันไปยังเฟรมใหม่ เมธอดสคิบ() จะถูกเรียกใช้ โดยจะคืนค่าหมายเลขเฟรมกลับมาซึ่งเวลาใหม่ของตัวเพลเออที่ทำงานอยู่จะถูกเปลี่ยนค่าตัวเลขไปด้วยเมธอด MediaTimeSetEvent()

เมธอดจะถูกเรียกใช้กับการเริ่มหรือการหยุดเล่นของตัวเพลเออ ซึ่งถ้าหากตัวเพลเอออยู่ในสถานะเริ่มต้น จะไม่สามารถระบุตำแหน่งเฟรมปัจจุบันได้อย่างถูกต้อง ส่งผลให้เฟรมที่จะข้ามไปนั้นไม่ถูกต้องตามไปด้วย หากตัวเพลเอออยู่ในสถานะเริ่มต้น เมธอดเซทเวลามีเดีย จะทำการเปลี่ยนสถานะของตัวเพลเออ และด้วยเมธอดตรวจสอบความเปลี่ยนแปลง() จะทำให้ตัวเพลเออทำงาน

ค่าตัวแปร : จำนวนของเฟรมที่จะสคิบ (framesToSkip) – จำนวนเฟรมที่ตัวเพลเออต้องการข้ามไปจากเฟรมปัจจุบัน หากมีค่าเป็นบวก ตัวเพลเออจะเล่นไปข้างหน้า (forward) แต่ถ้าหากมีค่าเป็นลบจะเป็นการเล่นย้อนกลับ (playback) เช่น สคิบ(-1) จะเป็นการย้อนกลับไปยังเฟรม 1 เฟรมก่อนหน้าเฟรมปัจจุบัน

ค่าที่ส่งกลับ : หมายเลขเฟรมที่ต้องการข้ามจากเฟรมปัจจุบันไปยังเฟรมที่ต้องการ

mapFrameToTime เปลี่ยนหมายเลขเฟรมเป็นเวลา

ค่าตัวแปร : frameNumber หมายเลขเฟรม

ค่าที่ส่งกลับ : เวลา หากกระบวนการล้มเหลวจะคืนค่าเป็น TIME UNKNOWN

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

mapTimeToFrame เปลี่ยนเวลาเป็นหมายเลขเฟรมจะคืนค่าหมายเลขเฟรมที่ตรงกับเวลาขณะนั้นหรือเป็นหมายเลขเฟรมที่ใกล้เคียงที่สุด

ค่าตัวแปร : mediaTime เวลาของมีเดีย

ค่าที่ส่งกลับ : หมายเลขเฟรมที่ตรงกับเวลาขณะนั้นหรือเป็นหมายเลขเฟรมที่ใกล้เคียงที่สุด หากกระบวนการล้มเหลวจะคืนค่าเป็น FRAME_UNKNOWN

2.4.2.2 การดักจับเฟรมจากวีดีโอสตรีม (FrameGrabbingControl)

เป็นอินเตอร์เฟสในการดักจับเฟรมจากวีดีโอสตรีม ซึ่งจะถูกส่งออกจากตัวเพลเออผ่านเมธอด getControl() ซึ่งเฟรมที่ถูกคืนค่ากลับมาจะเป็นรูปแบบการถอดรหัสที่ยังไม่ได้รับการปรับแต่ง ต้องอาศัยคลาสการแปลงรูป ImageConverter() ในการเปลี่ยนเป็นรูปแบบภาพของจาวา

Buffer	grabFrame() ดักจับเฟรมปัจจุบันจากวีดีโอสตรีม

grabFrame จับเฟรมปัจจุบันจากวีดีโอสตรีม เฟรมที่คืนค่าออกมานั้นจะอยู่ในรูปแบบของการถอดรหัสที่ไม่ได้รับการปรับแต่ง ต้องอาศัยคลาส การแปลงรูป ในการเปลี่ยนเข้าสู่รูปแบบภาพของจาวา ซึ่งหากตัวเพลเอออยู่ในสถานะเริ่มต้น จะไม่สามารถระบุเฟรมที่ต้องการคืนค่าออกมาได้อย่างถูกต้อง

ค่าที่ส่งกลับ : เฟรมที่คัดลอกจากเฟรมปัจจุบัน หากกระบวนการล้มเหลวจะคืนค่าเป็น null

2.4.2.3 การเปลี่ยนบัพเฟอร์ไปเป็นอิมเมจ (BufferToImage)

เป็นคลาสหลักในการเปลี่ยนวีดีโอ (video) บัพเฟอร์ เป็นเอดับบลิวทีอิมเมจ (AWT Image object) ที่สามารถจัดการได้ด้วยเมธอดของเอดับบลิวทีจะใช้คลาสนี้ร่วมกับการดักจับเฟรมจากวีดีโอสตรีม ในการดักจับเฟรมจากวีดีโอสตรีมและในการดำเนินการปรับแต่งภาพ

Image	BufferToImage(VideoFormat format) เปลี่ยน Buffer เป็น Image object ในรูปแบบที่ต้องการ
Java.awt.image	createImage(Buffer buffer) เปลี่ยน Buffer เป็นภาพในรูปแบบที่ต้องการเป็นมาตรฐานภาพแบบ AWT image และคืนค่ากลับมา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

BufferToImage เปลี่ยนบัพเฟอร์เป็นภาพในรูปแบบที่ต้องการเป็นมาตรฐาน

ค่าตัวแปร : format รูปแบบของ บัพเฟอร์

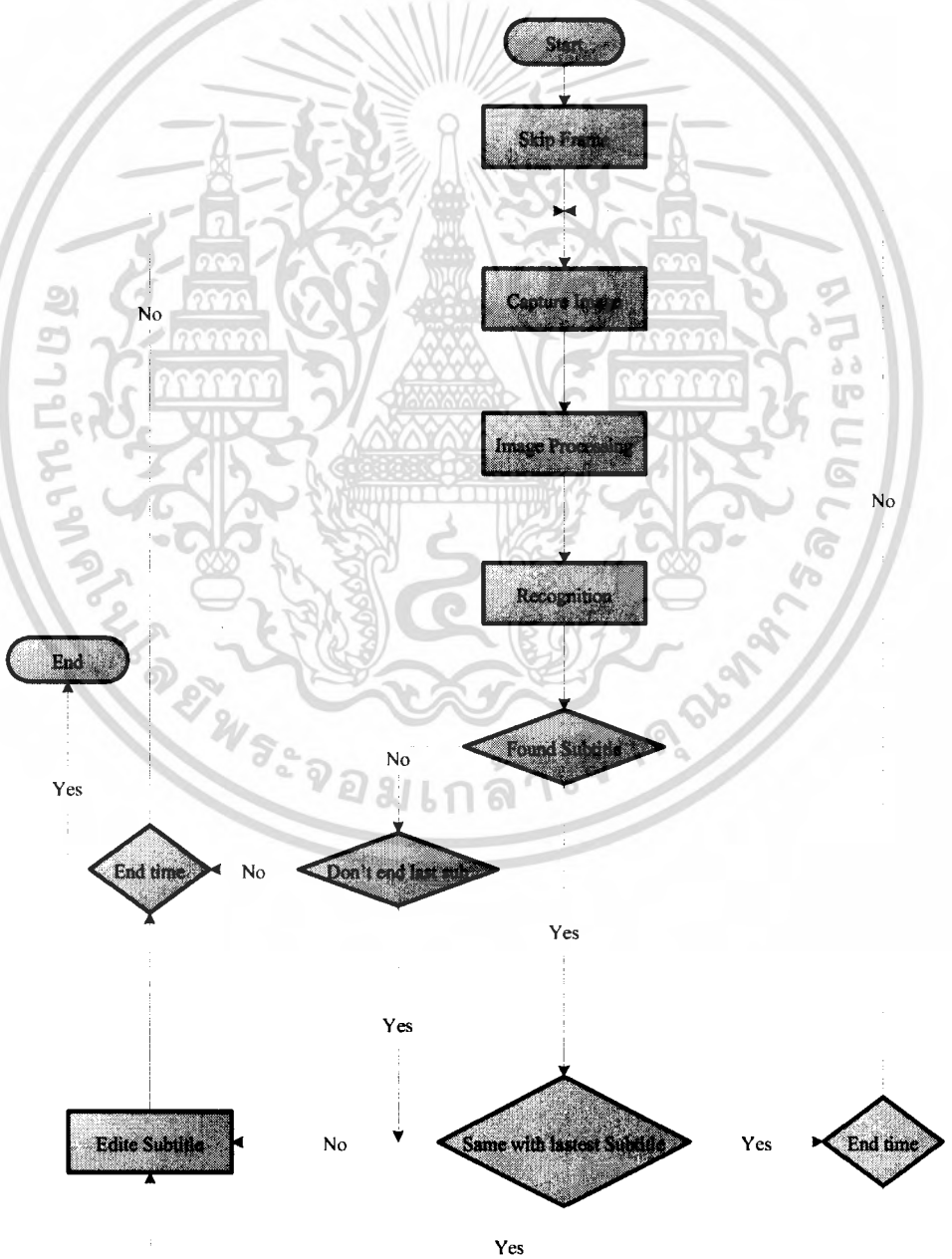
createImage เปลี่ยนบัพเฟอร์ที่ส่งเข้าไปเป็นมาตรฐานภาพแบบ AWT และส่งค่าคืนกลับมา ใน buffer จะประกอบด้วยข้อมูลวีดีโอในรูปแบบที่กำหนดไว้โดยผู้สร้าง หากค่าที่ส่งเข้ามาไม่สามารถประมวลผลได้จะคืนค่ากลับเป็น "null"

ค่าที่ส่งกลับ : ภาพเอดับบลิวที หากกระบวนการการแปลงล้มเหลวจะคืนค่ากลับเป็น "null"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3
การออกแบบโครงงาน

ในการออกแบบโครงสร้างของระบบเป็นการดำเนินการตั้งแต่การเริ่มเปิดภาพยนตร์เพื่อทำการดึงภาพจากเฟรมของภาพยนตร์ โดยจะนำภาพที่ได้มาทำประมวลผลภาพ เพื่อให้สะดวกกับรู้จำตัวอักษร จากนั้นจึงนำผลที่ได้มาสร้างคำบรรยายภาพยนตร์ที่มีการเก็บ ทั้งตัวอักษร และเวลา โดยจะมีการนำประโยคที่ได้มาทำการแปลภาษาอีกขั้นตอนหนึ่ง

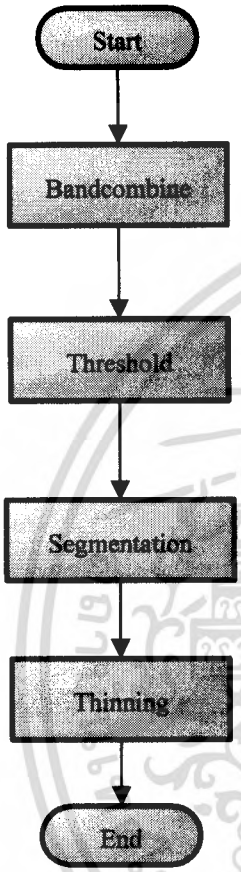


รูปที่ 3.1 Flow-chart การทำงานของขั้นตอนการทำงานหลัก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้ในการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนของส่วนการประมวลผลภาพ

จะแบ่งออกเป็นขั้นตอนย่อยๆ หลายกระบวนการโดยจะนำภาพที่ได้จากการ Capture Image แต่ละเฟรม มาทำการเปลี่ยนแปลงจากภาพมาตรฐาน RGB ด้วยกระบวนการต่างๆ ดังนี้



แบนก์คอมไบน์ (Bandcombine) เป็นขั้นตอนการแปลงจากภาพหลายๆ Band มาเป็น Bandเดียว เพื่อสามารถ Threshold

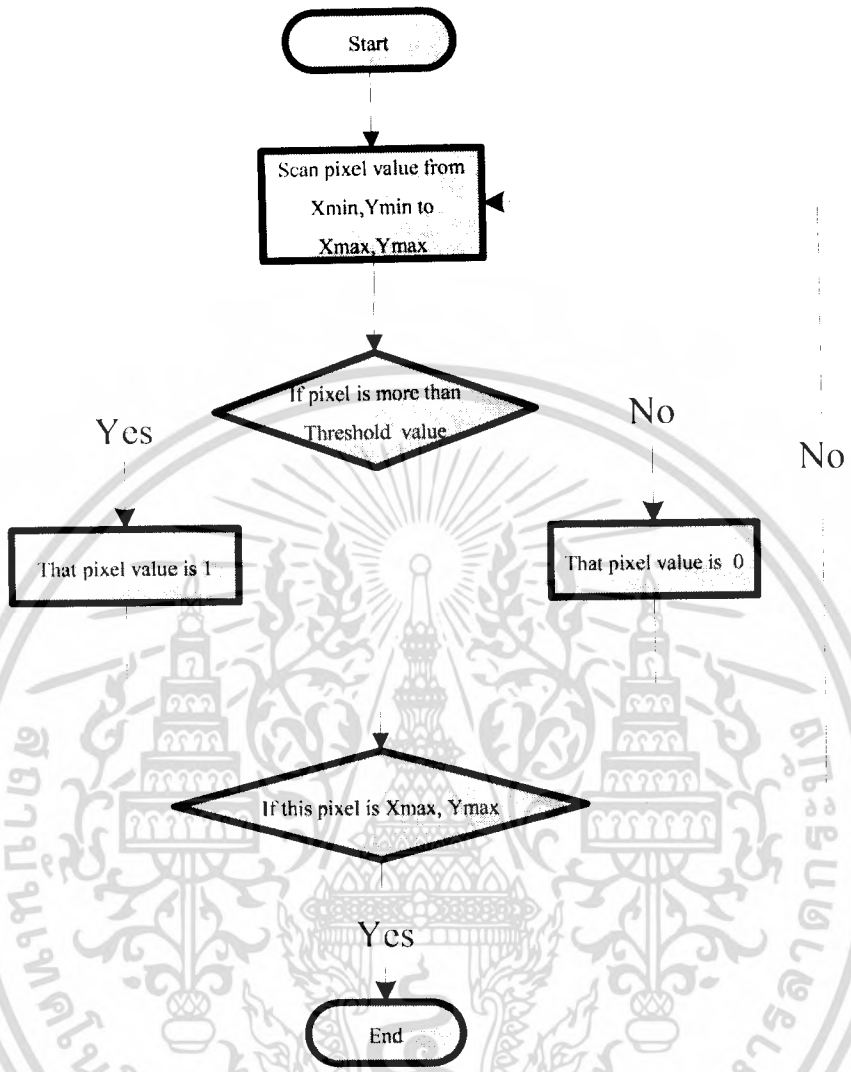
การทำเทรสโฮล (Threshold) โดยจะใช้มาตรฐาน 8 บิท ในการเก็บค่าซึ่ง สีดำคือ "0" และสีขาวคือ "255" (256 levels)

เซกเมนเทชัน (Segmentation) โดยจะทำการตัดตัวอักษรออกมาจากภาพ โดยตัดออกมาที่ละตัวเพื่อนำมาใช้ตรวจสอบตัวอักษร

การทำเส้นให้บาง (Thinning) เป็นการทำให้ความหนาตัวอักษรบางลงโดยใช้หลักการของการทำ Skeletonization Image เพื่อเป็นการหาแก่นภาพ เพื่อนำไปใช้ในการทำ Recognition ต่อไป

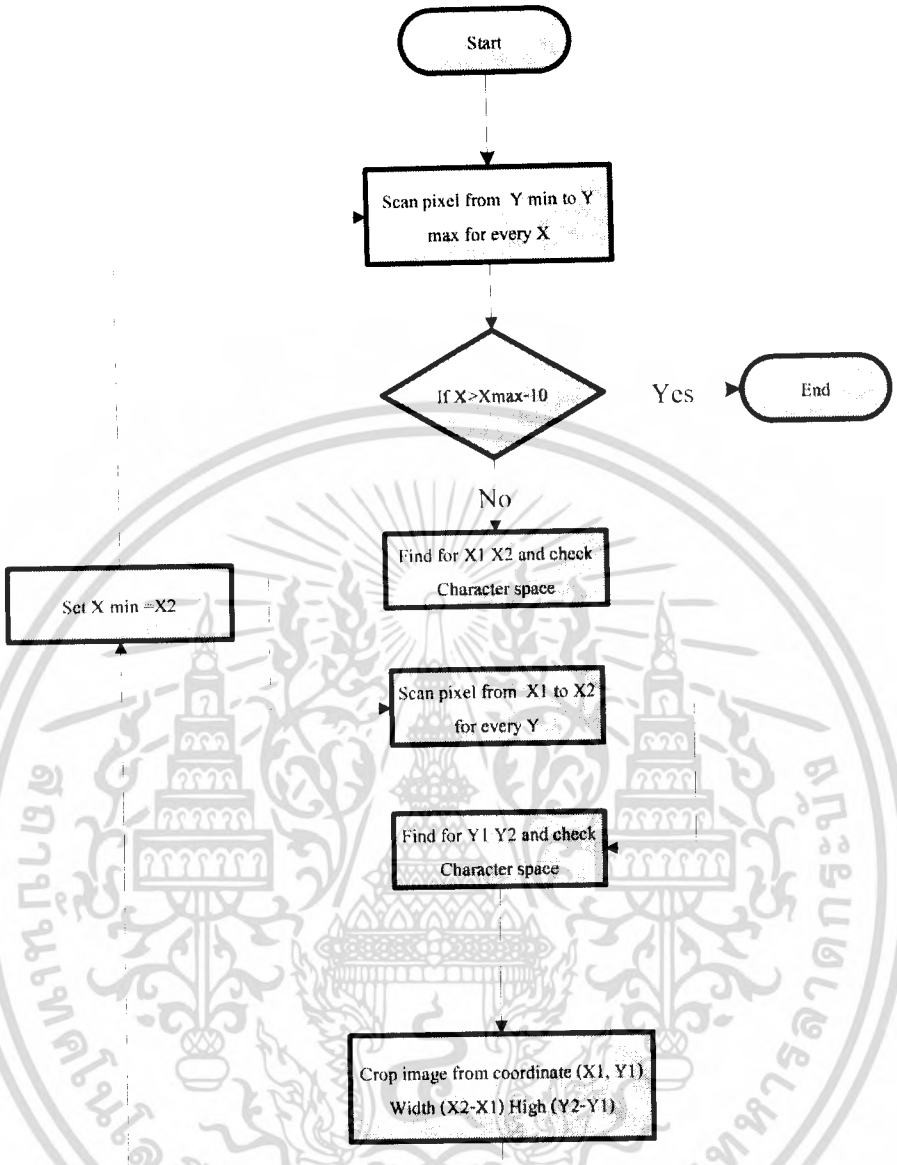
รูปที่ 3.2 Flow-chart การทำงานของ Image Processing

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



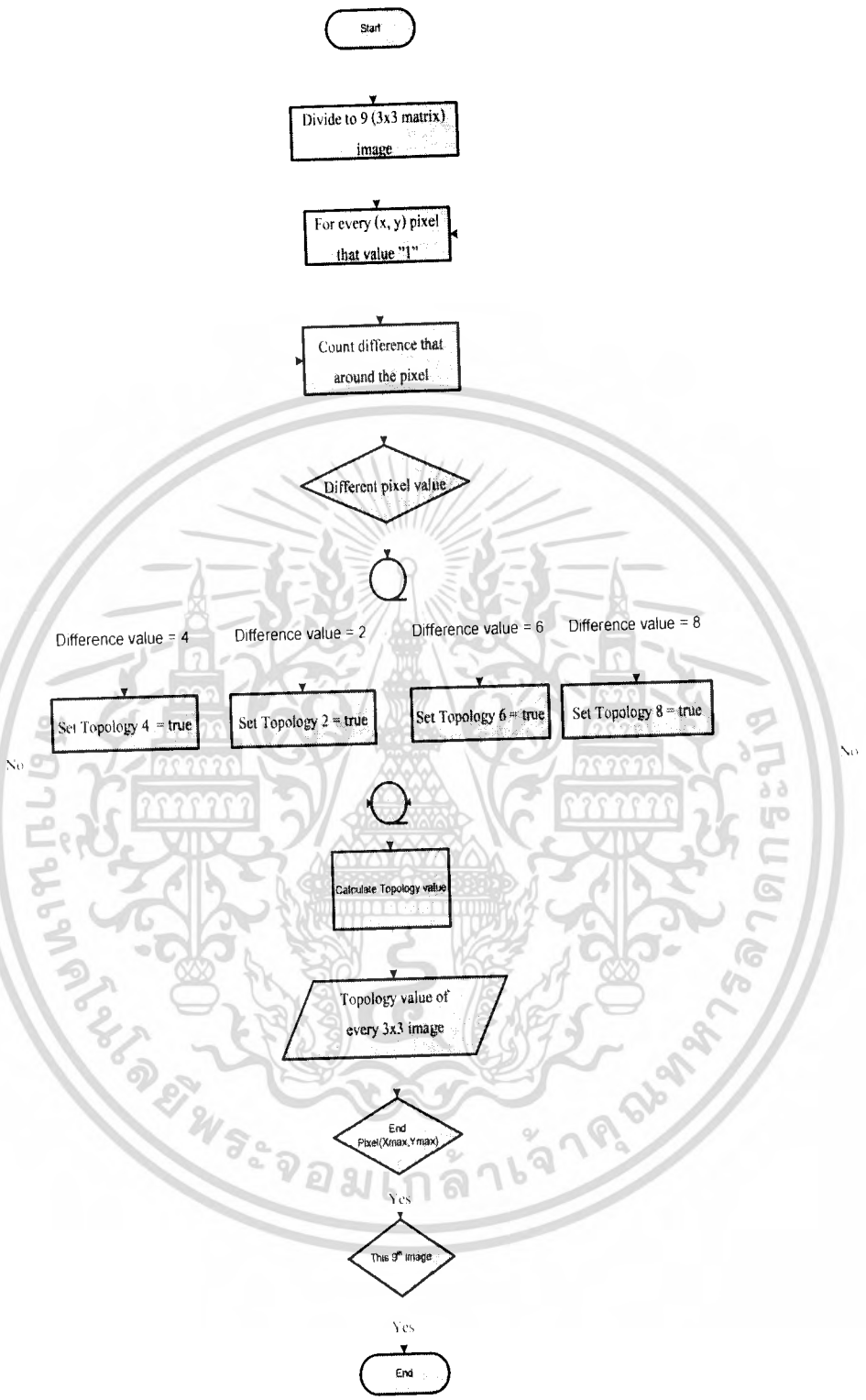
รูปที่ 3.3 Flow-chart การทำงานของ Threshold

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.4 Flow-chart การ Segment แต่ละตัวอักษร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.5 Flow-chart การคิดค่าTopology ของแต่ละตัวอักษรในกระบวนการ Recognition

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

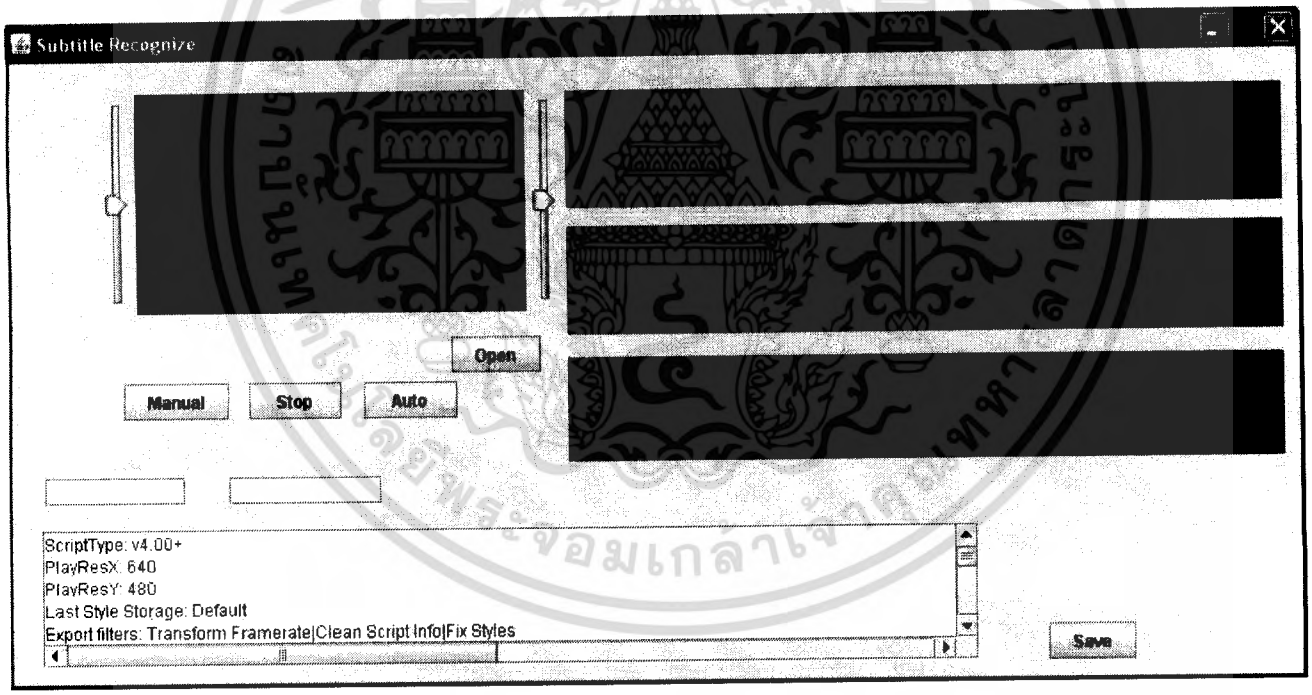
ผลการทดลอง

4.1 การทดลองการตรวจสอบคำบรรยายภาพยนตร์

การทดลองนี้จะทำการตรวจสอบโดยโปรแกรมที่ออกแบบโดยใช้ ภาษาจาวา (Java swing, JMF, JAI) ซึ่งโปรแกรมการจะทำเกิดความสะดวกในการใช้งาน โดยอุปกรณ์ที่ใช้ในการทดสอบมีรายการดังนี้

1. โปรแกรม ซับไทเทิลเรคคอกไนซ(Subtitle Recognize)
2. ภาพยนตร์ตัวอย่าง สกุลไฟล์ *.เอวีไอ (avi) มาตรฐานเอ็กซ์วิด (XVID)
3. โปรแกรมซับไทเทิลสตูดิโอ(Subtitle Studio) ที่ใช้ๆกันอยู่ทั่วๆไป

โดยการทดลองเป็นขั้นตอนการทำงานดังนี้

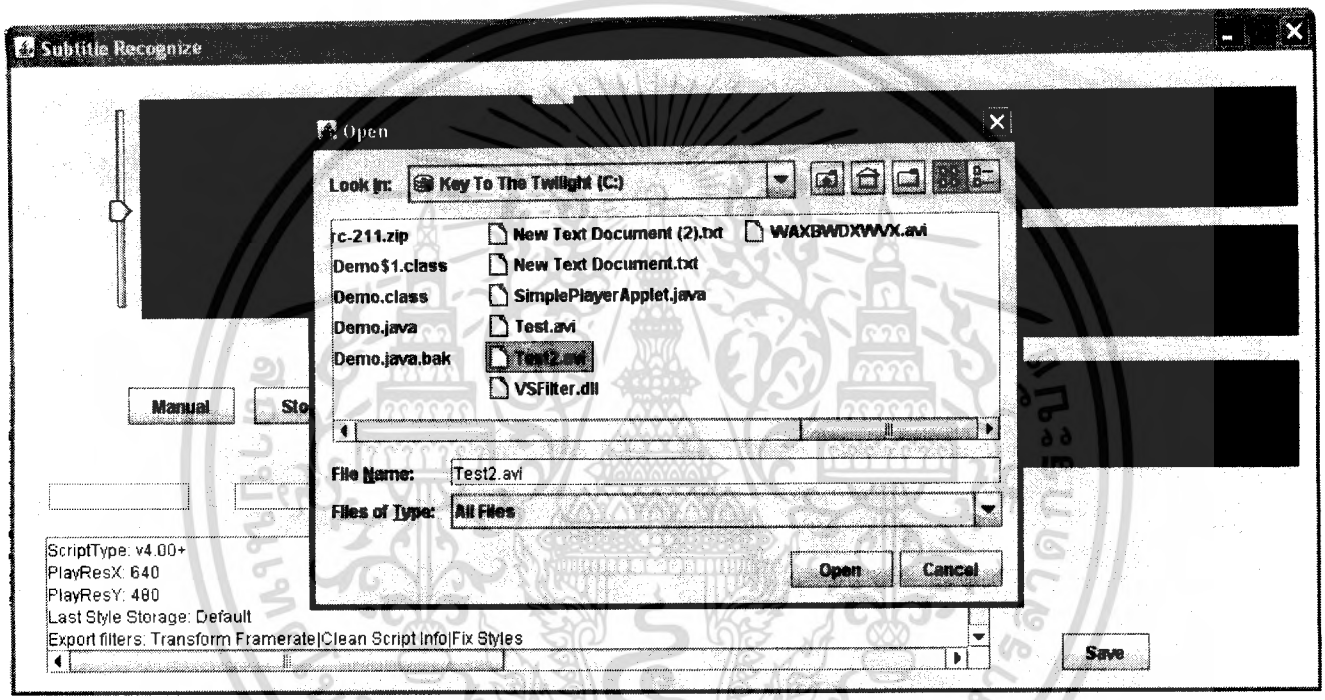


รูปที่ 4.1 หน้าจอแสดงการเริ่มต้นของโปรแกรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อเริ่มต้นการทำงานของโปรแกรมจะมีหน้าจอแสดงรายระเอียดขึ้นมาโดยเริ่มต้นยังไม่มีการกำหนดไฟล์ที่เรียกใช้ ดังนั้นจึงแสดงภาพเป็นสีดำทั้ง 4 อันซึ่งเราเรียกว่า ซับไตเติ้ลไลน์ (Subtitle line) เพราะค่าเป็น "null" และมีการกำหนดค่าเริ่มต้นของ ไฟล์ซับไตเติ้ล (subtitle files) ต้องเป็นมาตรฐานของสกุลไฟล์ .เอเอสเอสและสไลเดอร์ (slider) ทั้งสองข้างยังคงอยู่ที่ กึ่งกลาง

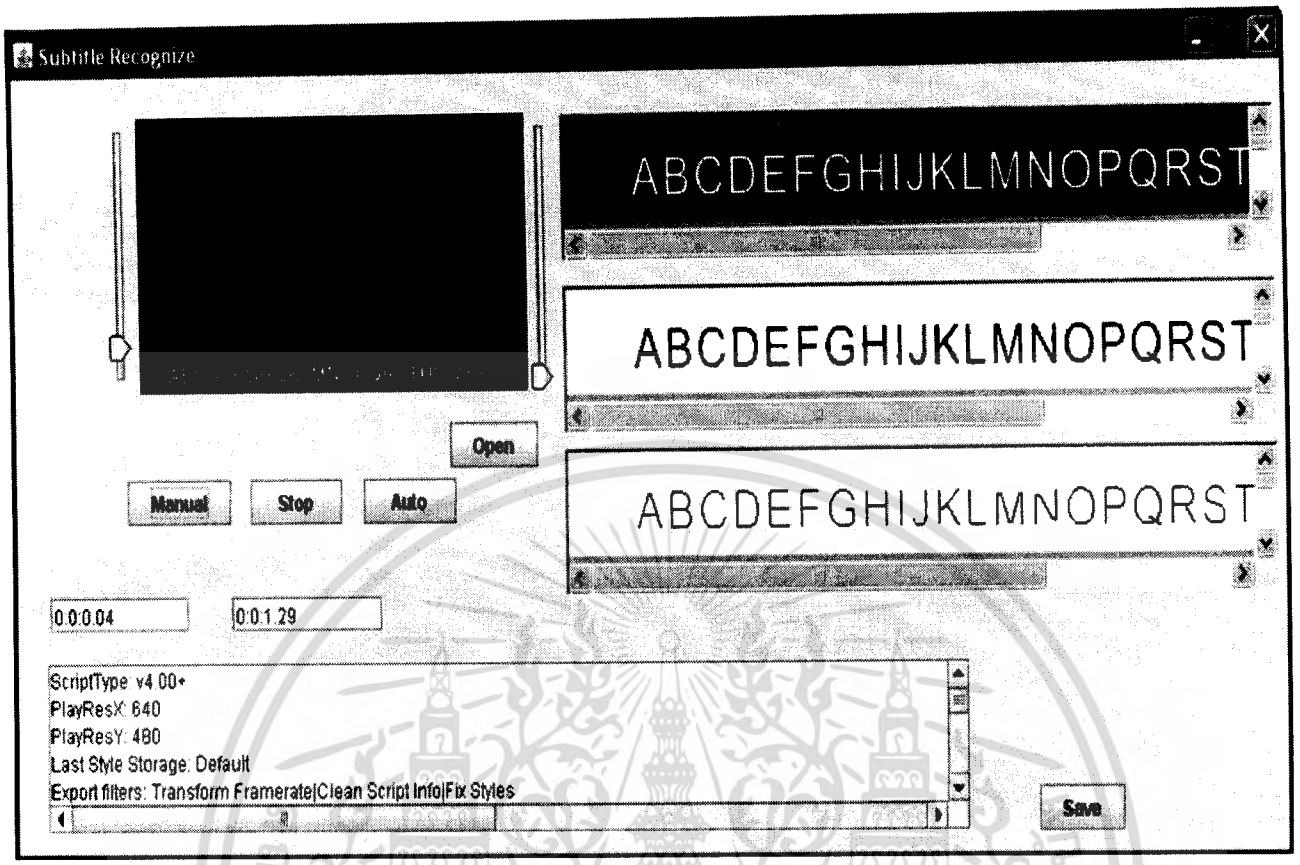
การเลือกทำงานของไฟล์ภาพยนตร์จะดำเนินการโดยรองรับ ภาพยนตร์แบบมาตรฐานสกุลไฟล์ .เอวีไอ โดยไม่สามารถรองรับการทำงานกับภาพยนตร์สกุลไฟล์อื่น



รูปที่ 4.2 หน้าจอการเลือกไฟล์เพื่อไปกำหนดเป็น Data Source ของโปรแกรม

โดยจะทำงานเลือกไฟล์จากไดเรกทอรี่ (Directory) แล้วแสดงผลออกมาที่หน้าต่างด้านซ้าย ซึ่งจะสามารถทำการเลือกขอบเขตการตรวจสอบ ของภาพยนตร์ที่แสดงได้ด้วยการเลื่อน สไลเดอร์ทั้ง 2 ที่อยู่ด้านข้างในการกำหนดขอบเขต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



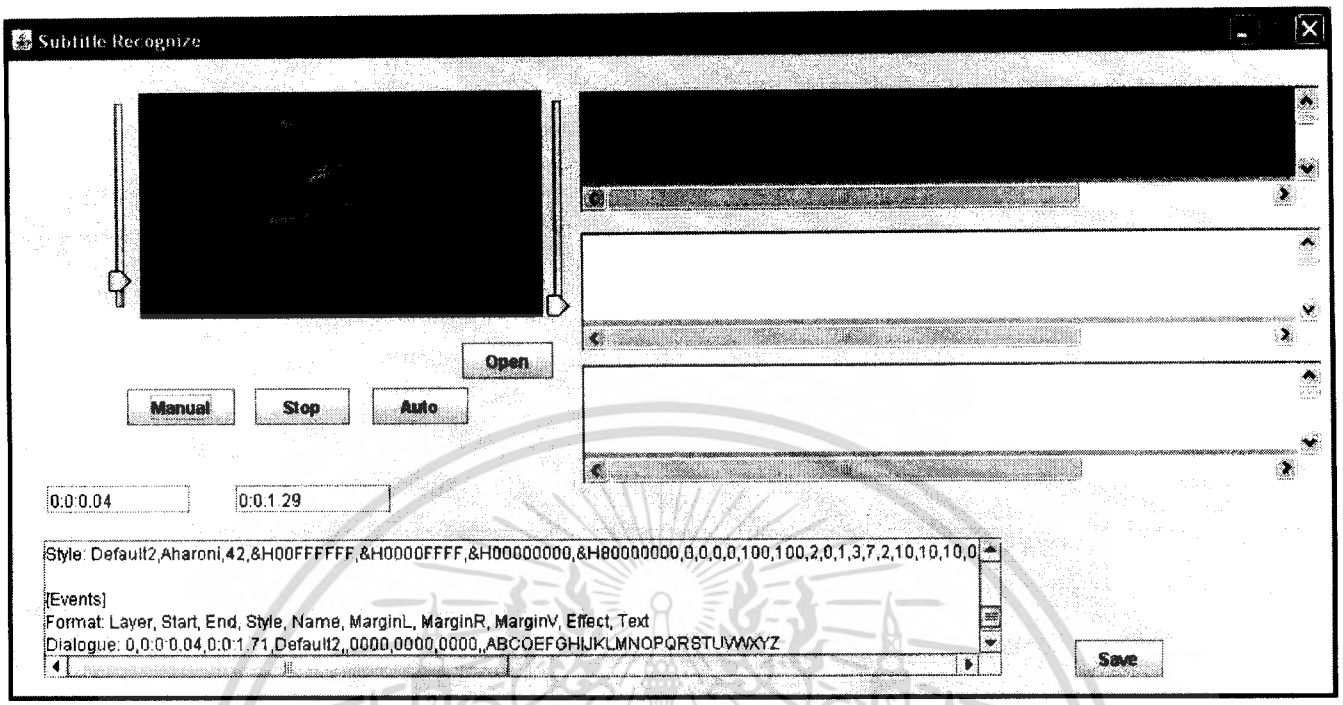
รูปที่ 4.3 หน้าจอการทำงานของโปรแกรม

ผลของการทดลอง โดยเมื่อทำการเลือกไฟล์โดยการกดปุ่ม "Open" แล้วทำการเลือกไฟล์เอวีไอขึ้นมา โปรแกรมจะเริ่มต้นทำงานโดยทำการแสดงภาพเริ่มต้นของภาพยนตร์มาแสดงในหน้าต่างด้านซ้าย แล้วแสดงส่วนที่ ตัดมาแสดงในหน้าต่างที่ 1 แล้วแสดงภาพที่ผ่านการแบนด์คอมไบน์, ไบนาไรซ์ (Binarize), อินเวอร์ส (Invert) ในหน้าต่างที่ 2 แล้วแสดงภาพที่ทำการเส้นให้บางในหน้าต่างที่ 3 โดยมีการแสดงเวลาที่เวลาเริ่มต้น และเวลาในขณะนั้นในแถบข้อความบนซับไตเติ้ลไลน์

โดยการที่โปรแกรมจะทำการอัพเดตซับไตเติ้ลไลน์ซึ่งจะเกิดขึ้นเมื่อมีเหตุการณ์ที่ตรงตามเงื่อนไขของโปรแกรมอย่างใดอย่างหนึ่ง

1. เมื่อมีการได้รับประโยคข้อความแล้วเกิดรับข้อความว่างเปล่าเข้ามา
2. เมื่อมีการได้รับประโยคข้อความแล้วเกิดรับข้อความที่ไม่เหมือนกับข้อความเก่าเข้ามา
3. เมื่อมีการได้รับประโยคข้อความแล้วเกิดแล้วภาพยนตร์สิ้นสุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



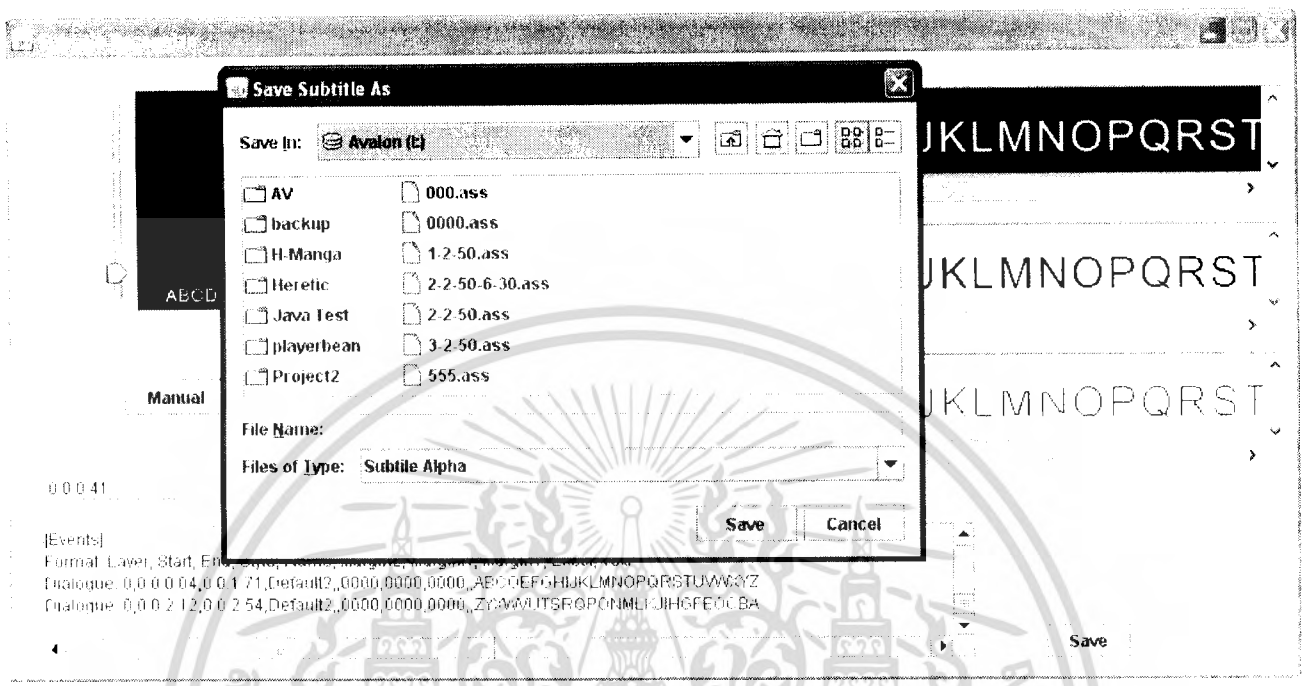
รูปที่ 4.4 หน้าจอผลการทำงานโดยการอัพเดทซับไตเติ้ลไลน์

การอัพเดตไลน์จะมีการอัพเดทตามกรณีที่ได้กล่าวมาโดยจะเก็บเวลาเริ่มต้นคำบรรยายภาพยนตร์ เวลาจบคำบรรยายภาพยนตร์ และคำบรรยายภาพยนตร์ที่เกิดขึ้น แล้วมีการอัพเดตในข้างท้ายของ Text Area ข้างล่าง โดยจะทำการอัพเดตลงไปในบรรทัดต่อๆ ไปโดยเก็บอยู่ในรูปแบบเดียวกันต่างกันในส่วนของเวลา และคำบรรยาย

โดยการทำงานสามารถใช้คำสั่งในการทำงานได้ 2 แบบ

1. "Manual" สั่งทำงานข้ามไปเป็นช่วงๆ ซึ่งจะมีการแสดงขั้นตอนการเก็บเวลาการเปลี่ยนแปลงของภาพทั้ง 4 หน้าต่าง แสดงขั้นตอนการ Update ที่ละช่วงๆ ตามที่เราสั่งหรือจะใช้คำสั่ง
2. "Auto" เพื่อที่จะให้โปรแกรมทำงานแบบอัตโนมัติ ซึ่งจะไม่มีการแสดงทางหน้าจอจนกว่า จะสิ้นสุดของภาพยนตร์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

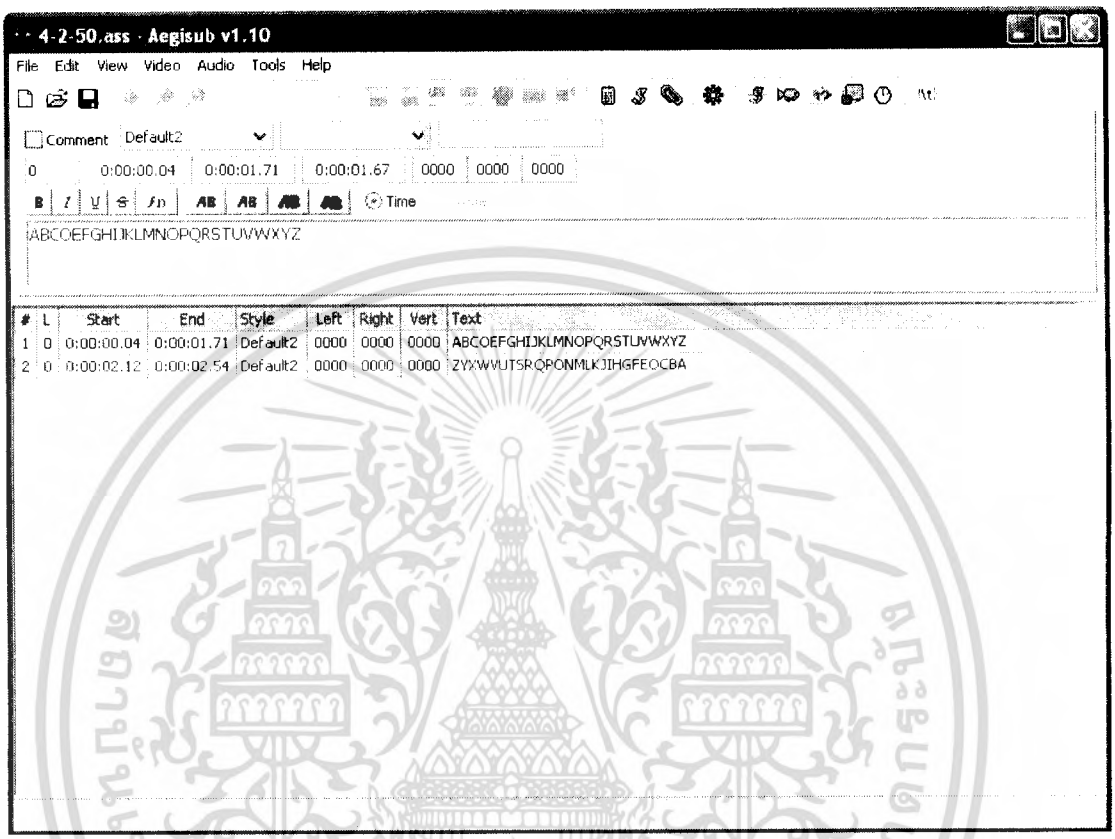


รูปที่ 4.5 หน้าจอการบันทึก Subtitle line ในนามสกุล .เอเอสเอสใน Directory ที่ต้องการ

เมื่อทำโปรแกรมถึงส่วนที่ต้องการแล้ว ด้วยวิธี “Manual” หรือ “Auto” ทำการซับไตเติ้ลไลน์ที่อยู่ในกล่องข้อความเป็นไฟล์ สกุล .เอเอสเอสซึ่งเป็นไฟล์ที่สามารถนำไปใช้ได้กับโปรแกรมซับไทเทิลสตูดิโอ ทั่วๆ ไปได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2 ผลการทดลองที่ได้กับโปรแกรมซับไทเทิลสตูดิโอทั่วไป



รูปที่ 4.6 หน้าจอผลการจดจำตัวอักษรที่ได้โดยโปรแกรม ซับไทเทิลสตูดิโอ Aegisub v1.10

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
สรุปผลการดำเนินงาน

การทดลองในโครงการนี้เป็นการจดจำตัวอักษรคำบรรยายภาพยนตร์ โดยใช้โปรแกรมซับไตเติ้ลเรคคอกไนซ์ (Subtitle Recognize) โดยภาษาจาวาซึ่งประสิทธิภาพที่ได้รับขึ้นอยู่กับลักษณะของภาพยนตร์ที่ในด้านความคมสว่างของภาพที่แสดง และรูปลักษณะของคำบรรยายภาพยนตร์ที่ใช้ว่า มีรูปแบบอักษรในลักษณะใด ซึ่งบางครั้งอาจทำให้เกิดการทำงานที่ผิดพลาดได้

5.1 ปัญหาที่พบในระหว่างการดำเนินโครงงาน

1. ในระบบการวิเคราะห์ตัวอักษรโดยใช้หลักการหาโทโปโลจีของตัวอักษรมีข้อบกพร่อง คือหากตัวอักษรมีลักษณะรูปร่างคล้ายๆ กัน จะเกิดความผิดพลาดขึ้นได้ในการเทียบตัวอักษร โดยไม่สามารถแยกออกได้ว่าเป็นตัวอักษรใด ทำให้อาจวิเคราะห์ผิดพลาด

2. ปัญหาเรื่องการแยกแยะตัวอักษรคำบรรยายภาพยนตร์ ออกจากภาพพื้นหลังนั้นจะมีความผิดพลาดหากสีของพื้นหลัง กับสีของตัวอักษรคำบรรยายภาพยนตร์ มีความใกล้เคียงกัน จะทำให้แยกแยะไม่ออก หรือเกิดความผิดพลาด โดยไม่สามารถแยกตัวอักษรที่จะไปนำวิเคราะห์ ออกมาเป็นส่วนๆ ได้ จึงทำให้ผลออกที่ได้เป็นว่างเปล่า หรือแยกตัวอักษรออกมาเป็นตัวที่ไม่ทราบได้

3. โดยทั่วไปภาพยนตร์มีรูปแบบตัวอักษรที่แตกต่างกัน โปรแกรมนี้เกิดข้อผิดพลาดหากลักษณะตัวอักษรมีความแตกต่างจากต้นแบบมากๆ โดยแต่ละรูปแบบตัวอักษรที่จะนำมาใช้ จะไม่สามารถใช้ร่วมกันได้หากมีความต่างกันมากจนไม่สามารถเทียบเคียงกันได้

5.2 แนวทางการแก้ไข

1. ทำการศึกษาค้นคว้าเกี่ยวกับการหาและแบ่งแยกโทโปโลยีเพื่อที่จะได้จำแนกหรือกำหนดลักษณะทางโทโปโลยีให้มากขึ้น และอาจนำคุณลักษณะของขนาดตัวอักษรที่เซกเมนเทชั่นได้มาเป็นตัวกำหนดเพิ่มเติม เพื่อช่วยในการวิเคราะห์ให้มีความมั่นยำมากขึ้น

2. อาจต้องทำการกำหนดคุณลักษณะของสีที่เลือกในการวิเคราะห์ให้มีความละเอียดให้มากขึ้นกว่านี้ และอาจนำเรื่องของการวิเคราะห์หาขอบภาพ การทำ Text area recognition เพื่อช่วยในการวิเคราะห์หาขอบเขตของคำบรรยาย และรูปแบบของอักษรได้

3. อาจต้องทำ library สำหรับการวิเคราะห์ตัวอักษรโดยแยกตามรูปแบบตัวอักษรให้มีความเพียงพอต่อความหลากหลายของรูปแบบอักษร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. อาจต้องทำ library สำหรับการวิเคราะห์ตัวอักษรโดยแยกตามรูปแบบตัวอักษรให้มีความเพียงพอต่อความหลากหลายของรูปแบบอักษร

5.3 แนวทางในการพัฒนา

แนวทางในการพัฒนาต่อ อาจแบ่งได้เป็น 2 แนวทาง

1. พัฒนาด้านแอปพลิเคชั่น ให้สามารถรองรับการทำงานกับไฟล์สกุลอื่นๆ ได้ไม่ให้จำเพาะเจาะจงที่ไฟล์รูปแบบสกุล .เอวีไอ สามารถเลือกรูปแบบของไฟล์คำบรรยายภาพยนตร์ที่ได้ให้มีความหลากหลายเพิ่มขึ้นเช่น .เอสเอสเอ, .เอสอาร์ที, .ซับ

2. พัฒนาด้าน User Interface ให้มีความสะดวกในการใช้งานให้มากขึ้นโดยอาจเพิ่มฟังชั่นในการทำงานให้มีความหลากหลาย หรือมีความรวดเร็วในการประมวลผลมากขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

[1] นิรันทร์ ทัศนลีลพร, สมศักดิ์ รัตนฤขัตร. 2539. **Printed Thai Character Recognitio** ปริญญานิพนธ์ ปริญญาวิศวกรรมศาสตร์ สาขาวิศวกรรมการควบคุมทางอุตสาหกรรม, สถาบันเทคโนโลยี พระจอมเกล้าเจ้าคุณทหารลาดกระบัง

[2] ดรวีระศักดิ์ ซึงถาวร. 2548. **JAVA Programming Volume** || กรุงเทพฯ: ซีเอดยูเคชั่น.

เวปไซต์อ้างอิง

http://ctsurin.no-ip.info/

http://en.wikipedia.org/wiki/

http://forum.java.sun.com/index.jspa

https://jaistuff.dev.java.net/

https://jai-demos.dev.java.net

http://yalor.yru.ac.th/

http://www.abstractnonsense.com/portfolio/ocr.html

http://www.cs.berkeley.edu/

http://www.geocities.com/yccheok81/zhangsuenthinning/index.html

http://www.ph.tn.tudelft.nl/

http://www.thaidev.com/

http://yalor.yru.ac.th/

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้