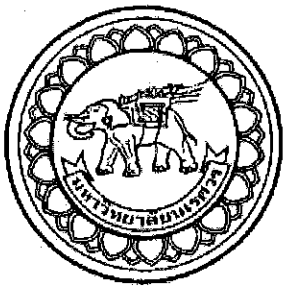
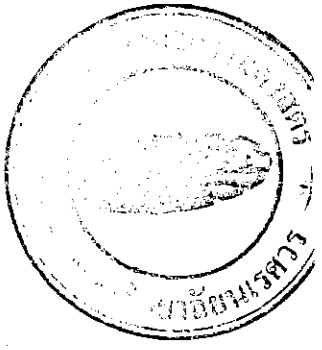




อัลกอริทึมสำหรับการตัดแบ่งวิดีโอแบบอัตโนมัติ

AN ALGORITHM FOR AUTOMATIC VIDEO SEGMENTATION

นางสาวนันทวดี โฉมยงค์ รหัส 48364777

นางสาวปวีณา สารชาติ รหัส 48364814

ห้องสมุด คณะวิศวกรรมศาสตร์
วันที่รับ 25 W.A. 2553
เลขทะเบียน 1497967
เลขเรียกหนังสือ ว.
จำนวนหน้า 94 ก 84 2551

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

สาขาวิชาวิศวกรรมคอมพิวเตอร์ ภาควิชาวิศวกรรมไฟฟ้าและคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร

ปีการศึกษา 2551

ใบรับรองโครงงานวิศวกรรม

หัวข้อโครงงาน อัลกอริทึมสำหรับการตัดแบ่งวิดีโอแบบอัตโนมัติ

ผู้ดำเนินโครงงาน **นางสาวนันทวดี โฉมยงค์ รหัส 48364777**

นางสาวปวีณา สารชาติ รหัส 48364814

อาจารย์ที่ปรึกษา ดร.ไพศาล มุณีสว่าง

สาขาวิชา วิศวกรรมคอมพิวเตอร์

ภาควิชา วิศวกรรมไฟฟ้าและคอมพิวเตอร์

ปีการศึกษา 2551

.....

คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร อนุมัติให้โครงงานฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิศวกรรมศาสตรบัณฑิต สาขาวิชาวิศวกรรมคอมพิวเตอร์

คณะกรรมการสอบโครงงานวิศวกรรม

..... ประธานกรรมการ

(ดร.ไพศาล มุณีสว่าง)

..... กรรมการ

(ดร.สุรเดช จิตประไพกุลศาล)

..... กรรมการ

(ดร.พนมขวัญ ริยะมงคล)

หัวข้อโครงงาน	อัลกอริทึมสำหรับการตัดแบ่งวิดีโอแบบอัตโนมัติ		
ผู้ดำเนินโครงงาน	นางสาวนันทวดี	โฉมยงค์	รหัส 48364777
	นางสาวปวีณา	สารชาติ	รหัส 48364814
อาจารย์ที่ปรึกษา	ดร.ไพศาล มุณีสว่าง		
สาขาวิชา	วิศวกรรมคอมพิวเตอร์		
ภาควิชา	วิศวกรรมไฟฟ้าและคอมพิวเตอร์		
ปีการศึกษา	2551		

บทคัดย่อ

โครงงานนี้เป็นการศึกษาและพัฒนาโปรแกรมสำหรับการตัดแบ่งวิดีโอแบบอัตโนมัติ โดยวิดีโอที่ใช้ในโครงงานเป็นวิดีโอข่าวที่มีความต่อเนื่องกัน 2-3 ข่าวขึ้นไป ซึ่งผู้จัดทำใช้คุณสมบัติ Color Descriptor ของ MPEG-7 ในการอธิบายคุณลักษณะของสีของเฟรมภาพ โดย Descriptor ที่เลือกใช้คือ Color Layout โปรแกรมนี้แบ่งออกเป็น 2 ขั้นตอน คือ การจับคู่เปรียบเทียบ (Matching) ระหว่างเฟรมภาพของวิดีโอข่าวที่อยู่ติดกัน เพื่อทำการเปรียบเทียบค่าความแตกต่าง (Distance) ระหว่างกัน หากเฟรมภาพคู่ใดมีค่าความแตกต่างต่อกันมาก ก็อาจจะอนุมานได้ว่าเป็นเฟรมภาพจากคนละเนื้อหาข่าว จึงทำการตัดวิดีโอระดับช็อต ณ ตำแหน่งนั้น ๆ อีกขั้นตอนหนึ่งคือทำการเลือกคีย์เฟรมที่เป็นผู้ประกาศข่าวมาทำการเปรียบเทียบกับช็อตต่าง ๆ ที่ตัดได้อีกครั้ง เพื่อให้มีความแม่นยำมากกว่าการตัดเพียงครั้งเดียว หากพบจุดที่เฟรมใด ๆ มีค่าความแตกต่างกับคีย์เฟรมน้อยมาก และมีค่าต่ำกว่าค่าเทรสโฮลด์ที่กำหนด ให้ทำการตัดแบ่งในระดับสตอรี่ ซึ่งจะได้ผลลัพธ์มาเป็นไฟล์วิดีโอข่าวย่อยที่มีเนื้อหาข่าว 1 ข่าวต่อ 1 ไฟล์วิดีโอ โดยเครื่องมือที่ใช้ในการเขียนโปรแกรมของโครงงานนี้ คือ Microsoft Visual C# .Net

Project Title	AN ALGORITHM FOR AUTOMATIC VIDEO SEGMENTATION
Name	Miss Nanthawadee Chomyong ID. 48364777
	Miss Paweena Sarachart ID. 48364814
Project Advisor	Paisarn Muneesawang, Ph.D.
Major	Computer Engineering.
Department	Electrical and Computer Engineering.
Academic Year	2551

ABSTRACT

This project is to study and develop a program that can do automatic video segmentation. The condition of the input video stream is to be at least two or three continuous news videos. This project uses the color descriptor of MPEG-7 to differentiate the frame colors from the stream. There are two parts in this program : The first part is to choose the color layout descriptor to match two parts of the closest frame, and to find the differences between them. If any pair of the frames has many differences between each other, then it is segmented from that position onwards. The next part is to choose the frame that has the news announcer and compare that with all the shots from the first part. If it is found to have differences less than the set threshold, the video will be segmented here to produce one story. The output will be a single video file that contains one news per file. The tools used to complete this project was Microsoft Visual C# .Net .

กิตติกรรมประกาศ

โครงงานนี้จะสำเร็จลุล่วงมิได้ หากไม่มี ดร.ไพศาล มุณีสว่าง ที่ได้ให้คำปรึกษาและคำแนะนำในทุกขั้นตอนของการจัดทำโครงงานด้วยดีเสมอมา รวมถึงเสียสละเวลาเพื่อช่วยตรวจสอบและเสนอแนะข้อบกพร่องที่ควรแก้ไขของโปรแกรมในขั้นตอนต่าง ๆ เสมอ

ดร.สุรเดช จิตประไพกุลศาล ที่ได้ให้คำแนะนำในการวางแผนการออกแบบและพัฒนาโปรแกรมอย่างเป็นขั้นตอน ช่วยให้มีความรอบคอบและมีระบบระเบียบในการเขียนโปรแกรม รวมถึง ดร.พนมขวัญ ริยะมงคล ที่ให้คำแนะนำในการปรับปรุงโครงงานให้ดียิ่งขึ้น จนทำให้การจัดทำโครงงานนี้สำเร็จลุล่วงไปได้ด้วยดี

ทั้งนี้ ผู้พัฒนาโครงงานขอขอบพระคุณคุณพ่อ คุณแม่และครอบครัว ที่คอยให้กำลังใจตลอดเวลา เป็นหนึ่งในเบื้องหลังของความสำเร็จในการพัฒนาโครงงานนี้, เพื่อน ๆ ที่ช่วยเป็นกำลังใจและให้ความช่วยเหลือเสมอ รวมถึงคณาจารย์ทุกท่านที่ได้ให้ความรู้ต่าง ๆ ตลอดมา

ผู้พัฒนาโครงงานจึงขอขอบพระคุณทุกท่านไว้ ณ โอกาสนี้

นางสาวนันทวดี โฉมยงค์

นางสาวปวีณา สารชาติ

สารบัญ

หน้า

บทคัดย่อภาษาไทย .....ก

บทคัดย่อภาษาอังกฤษ .....ข

กิตติกรรมประกาศ .....ค

สารบัญ .....ง

สารบัญตาราง .....ช

สารบัญรูป .....ซ

บทที่ 1 บทนำ

1.1 ที่มาและความสำคัญ 1

1.2 วัตถุประสงค์ของโครงการ 2

1.3 ขอบข่ายของโครงงาน 2

1.4 ขั้นตอนการดำเนินงาน 2

1.5 แผนการดำเนินงาน 3

1.6 รายละเอียดงบประมาณของโครงงาน 3

บทที่ 2 ทฤษฎีที่เกี่ยวข้อง

2.1 การตัดแบ่งวิดีโอ (Video Segmentation) 4

2.1.1 ความรู้พื้นฐานเกี่ยวกับวิดีโอ 4

2.1.1.1 คุณภาพของวิดีโอ 5

2.1.1.2 มาตรฐานของภาพวิดีโอ 5

2.1.1.3 การบีบอัดวิดีโอ 6

2.1.2 ความรู้พื้นฐานเกี่ยวกับการตัดแบ่งวิดีโอ (Video Segmentation) 6

2.1.2.1 องค์ประกอบของภาพวิดีโอ 6

2.2 ระบบสี Y/Cb/Cr 11

2.3 MPEG-7 11

2.3.1 ทฤษฎี Color Layout Descriptor ของ MPEG-7 12

2.3.1.1 การดึงข้อมูลจากภาพในระบบสี Y/Cb/Cr (Extraction) 12

2.3.1.2 การจับคู่เปรียบเทียบ (Matching) 14

สารบัญ (ต่อ)

หน้า

2.3.1.3 ตัวอย่างการคำนวณ Color Layout ของภาพตัวอย่าง 15

บทที่ 3 การออกแบบระบบ

3.1 เป้าหมายของการออกแบบโปรแกรม 16

3.2 ลักษณะของข้อมูลเข้า (Input) และข้อมูลออก (Output) 16

3.3 ผังงาน (Flowchart) 18

3.4 วิธีดำเนินการ 19

3.5 การวิเคราะห์ความถูกต้อง 25

3.6 เครื่องมือที่ใช้พัฒนาโครงงาน 26

3.7 การออกแบบหน้าตาของโปรแกรม 26

3.8 การออกแบบของโปรแกรม 27

3.8.1 Use Case Diagram 27

3.8.2 Class Diagram 28

3.8.3 Sequence Diagram 29

3.8.4 Activity Diagram 30

บทที่ 4 การดำเนินงานและผลการทดลอง

4.1 ข้อมูลเบื้องต้นของไฟล์วิดีโอข่าว 32

4.2 ผลการทดลองตัดแบ่งวิดีโอข่าว 32

4.3 การทำงานของโปรแกรม 54

บทที่ 5 บทสรุป

5.1 บทสรุป 59

5.2 สรุปขั้นตอนการทำงาน 59

5.3 ปัญหาและอุปสรรคในการพัฒนา 60

5.4 ข้อเสนอแนะ 60

5.5 แนวทางในการพัฒนาต่อในอนาคต 60

5.6 สรุปผลการดำเนินการ 60

สารบัญ (ต่อ)

หน้า

เอกสารอ้างอิง 61

ประวัติผู้เขียนโครงงาน 62

สารบัญตาราง

ตารางที่		หน้า
1.1	แผนการดำเนินงาน	3
2.1	ตารางแสดงรายละเอียดข้อมูลสัมประสิทธิ์	14
4.1	รายละเอียดของไฟล์วิดีโอข่าว	32
4.2	ผลลัพธ์การตัดแบ่งวิดีโอในเทรสโฮลด์ระดับต่างๆ ของไฟล์วิดีโอ news_nbt_01	35
4.3	ผลลัพธ์การตัดแบ่งวิดีโอในเทรสโฮลด์ระดับต่างๆ ของไฟล์วิดีโอ news_nbt_02	36
4.4	ผลลัพธ์การตัดแบ่งวิดีโอในเทรสโฮลด์ระดับต่างๆ ของไฟล์วิดีโอ news_ch3_01	37
4.5	ผลลัพธ์การตัดแบ่งวิดีโอในเทรสโฮลด์ระดับต่างๆ ของไฟล์วิดีโอ news_ch3_02	39
4.6	ผลลัพธ์การตัดแบ่งวิดีโอในเทรสโฮลด์ระดับต่างๆ ของไฟล์วิดีโอ news_ch3_03	40
4.7	ผลลัพธ์การตัดแบ่งวิดีโอในเทรสโฮลด์ระดับต่างๆ ของไฟล์วิดีโอ news_modern-9_01	42
4.8	ผลลัพธ์การตัดแบ่งวิดีโอในเทรสโฮลด์ระดับต่างๆ ของไฟล์วิดีโอ news_modern-9_02	43
4.9	ผลลัพธ์การตัดแบ่งวิดีโอในเทรสโฮลด์ระดับต่างๆ ของไฟล์วิดีโอ news_modern-9_02	44
4.10	ผลลัพธ์การตัดแบ่งวิดีโอในเทรสโฮลด์ระดับต่างๆ ของไฟล์วิดีโอ news_modern-9_03	46
4.11	ผลลัพธ์การตัดแบ่งวิดีโอในเทรสโฮลด์ระดับต่างๆ ของไฟล์วิดีโอ news_modern-9_04	47
4.12	ค่าความถูกต้องเมื่อให้ค่าเทรสโฮลด์เท่ากับ 5	48
4.13	ค่าความถูกต้องเมื่อให้ค่าเทรสโฮลด์เท่ากับ 10	48
4.14	ค่าความถูกต้องเมื่อให้ค่าเทรสโฮลด์เท่ากับ 15	49
4.15	เปรียบเทียบผลลัพธ์ของโปรแกรม	49
4.16	ค่าความผิดพลาดเมื่อวัดโดย False Positive และ False Negative	51
4.17	ระยะเวลาการประมวลผลของโปรแกรม	52
4.18	ค่าความถูกต้องเมื่อกำหนดให้ผู้ใช้สามารถเลือกเฟรมได้ 3 เฟรม และให้ค่าเทรสโฮลด์เท่ากับ 15	53

สารบัญรูป

รูปที่		หน้า
2.1	องค์ประกอบของภาพวิดีโอ	6
2.2	สตอรี่ของวิดีโอข่าว	7
2.3	ซีนของวิดีโอข่าว	7
2.4	ช็อตของวิดีโอข่าว	8
2.5	เทคนิคการเชื่อมต่อช็อตแบบคัต (Cut)	9
2.6	เทคนิคการเชื่อมต่อช็อตแบบเฟดอิน (Fade In)	9
2.7	เทคนิคการเชื่อมต่อช็อตแบบเฟดเอาต์ (Fade Out)	10
2.8	เทคนิคการเชื่อมต่อช็อตแบบดิสโซลว์ (Dissolve)	10
2.9	เทคนิคการเชื่อมต่อช็อตแบบไวป์ (Wipe)	11
2.10	กระบวนการของ Descriptor ในการดึงข้อมูลจากภาพ	13
2.11	การสแกนแบบ Zig-zag	14
2.12	ภาพคู่ตัวอย่างที่ 1 ในการคำนวณค่า Color Layout	15
2.13	ภาพคู่ตัวอย่างที่ 2 ในการคำนวณค่า Color Layout	15
3.1	โครงสร้างของวิดีโอข่าว	16
3.2	ลักษณะของอินพุทและเอาท์พุทของระบบ	17
3.3	ลักษณะการทำงานโดยรวมของระบบ	17
3.4	ขั้นตอนการดำเนินการของโปรแกรม	18
3.5	ไฟล์วิดีโอตัวอย่างในการดำเนินการตามโปรแกรม	19
3.6	เฟรมที่เลือกมาเป็นคีย์เฟรมในระดับช็อต	20
3.7	เฟรมที่เลือกมาเรียงกันตามเวลา	20
3.8	กราฟแสดงความแตกต่างระหว่างเฟรมที่ติดกัน	21
3.7	คีย์เฟรมที่ได้จากการตัดแบ่งระดับช็อต	21
3.8	กราฟแสดงความแตกต่างระหว่างคีย์เฟรมที่ได้	22
3.9	กราฟแสดงความแตกต่างระหว่างคีย์เฟรมที่ 1 กับเฟรมที่ได้จากการตัดในระดับช็อต	23
3.10	กราฟแสดงความแตกต่างระหว่างคีย์เฟรมที่ 5 กับเฟรมที่ได้จากการตัดในระดับช็อต	23
3.11	กราฟแสดงความแตกต่างระหว่างคีย์เฟรมที่ 7 กับเฟรมที่ได้จากการตัดในระดับช็อต	24
3.12	ผลลัพธ์ที่ได้จากการตัดแบ่งไฟล์ภาพวิดีโอในขั้นสุดท้าย	24

สารบัญรูป (ต่อ)

รูปที่		หน้า
3.13	ไฟล์ภาพวิดีโอ 5 เรื่องจากการตัดแบ่ง	25
3.14	ลักษณะรูปแบบของโปรแกรมที่ออกแบบ	26
3.15	Use Case Diagram	27
3.16	Class Diagram สำหรับให้ผู้ใช้เลือก 1 เฟรมในการเปรียบเทียบระดับสตอรี่	28
3.17	Class Diagram สำหรับให้ผู้ใช้เลือก 3 เฟรมในการเปรียบเทียบระดับสตอรี่	28
3.18	Sequence Diagram	29
3.19	Activity Diagram สำหรับให้ผู้ใช้เลือก 1 เฟรมในการเปรียบเทียบระดับสตอรี่	30
3.20	Activity Diagram สำหรับให้ผู้ใช้เลือก 3 เฟรมในการเปรียบเทียบระดับสตอรี่	31
4.1	กราฟความสัมพันธ์ระหว่างเฟรมกับค่าความแตกต่างของไฟล์วิดีโอ news_ch3_01	33
4.2	กราฟความสัมพันธ์ระหว่างเฟรม กับค่าความแตกต่างของไฟล์วิดีโอ news_modern-9_01	33
4.3	กราฟความสัมพันธ์ระหว่างเฟรมกับค่าความแตกต่างของไฟล์วิดีโอ news_nbt_01	33
4.4	ตัวอย่างภาพในไฟล์วิดีโอข่าว news_nbt_01	34
4.5	กราฟความสัมพันธ์ระหว่างคีย์เฟรมของผู้ประกาศข่าว กับคีย์เฟรมที่ตัดในระดับช็อตของไฟล์วิดีโอ news_nbt_01	35
4.6	ตัวอย่างภาพในไฟล์วิดีโอข่าว news_nbt_02	36
4.7	กราฟความสัมพันธ์ระหว่างคีย์เฟรมของผู้ประกาศข่าว กับคีย์เฟรมที่ตัดในระดับช็อตของไฟล์วิดีโอ news_nbt_02	36
4.8	ตัวอย่างภาพในไฟล์วิดีโอข่าว news_ch3_01	37
4.9	กราฟความสัมพันธ์ระหว่างคีย์เฟรมของผู้ประกาศข่าว กับคีย์เฟรมที่ตัดในระดับช็อตของไฟล์วิดีโอ news_ch3_01	38
4.10	ตัวอย่างภาพในไฟล์วิดีโอข่าว news_ch3_02	38
4.11	กราฟความสัมพันธ์ระหว่างคีย์เฟรมของผู้ประกาศข่าว กับคีย์เฟรมที่ตัดในระดับช็อตของไฟล์วิดีโอ news_ch3_02	32
4.12	ตัวอย่างภาพในไฟล์วิดีโอข่าว news_ch3_03	40
4.13	กราฟความสัมพันธ์ระหว่างคีย์เฟรมของผู้ประกาศข่าว กับคีย์เฟรมที่ตัดในระดับช็อตของไฟล์วิดีโอ news_ch3_03	41
4.14	ตัวอย่างภาพในไฟล์วิดีโอข่าว news_modern-9_01	41

สารบัญรูป (ต่อ)

รูปที่		หน้า
4.15	กราฟความสัมพันธ์ระหว่างคีย์เฟรมของผู้ประกาศข่าว กับคีย์เฟรมที่ตัดในระดับช็อตของไฟล์วิดีโอ news_modern-9_01	42
4.16	ตัวอย่างภาพในไฟล์วิดีโอข่าว news_modern-9_02	43
4.17	กราฟความสัมพันธ์ระหว่างคีย์เฟรมของผู้ประกาศข่าวคนที่หนึ่ง กับคีย์เฟรมที่ตัดในระดับช็อตของไฟล์วิดีโอ news_modern-9_02	43
4.18	ตัวอย่างภาพในไฟล์วิดีโอข่าว news_modern-9_02	44
4.19	กราฟความสัมพันธ์ระหว่างคีย์เฟรมของผู้ประกาศข่าวคนที่สอง กับคีย์เฟรมที่ตัดในระดับช็อตของไฟล์วิดีโอ news_modern-9_02	45
4.20	ตัวอย่างภาพในไฟล์วิดีโอข่าว news_modern-9_03	45
4.21	กราฟความสัมพันธ์ระหว่างคีย์เฟรมของผู้ประกาศข่าว กับคีย์เฟรมที่ตัดในระดับช็อตของไฟล์วิดีโอ news_modern-9_03	46
4.22	ตัวอย่างภาพในไฟล์วิดีโอข่าว news_modern-9_04	47
4.23	กราฟความสัมพันธ์ระหว่างคีย์เฟรมของผู้ประกาศข่าว กับคีย์เฟรมที่ตัดในระดับช็อตของไฟล์วิดีโอ news_modern-9_04	47
4.24	ผลลัพธ์ของการตัดแบ่งไฟล์วิดีโอ news_modern-9_04	50
4.25	ผลลัพธ์ของการตัดแบ่งไฟล์วิดีโอ news_modern-9_01	51
4.26	ผลลัพธ์ของการตัดแบ่งไฟล์วิดีโอ news_modern-9_01 เมื่อให้ผู้ใช้เลือกได้ 3 เฟรม	54
4.27	หน้าตาของโปรแกรม	55
4.28	โปรแกรมขณะรับไฟล์วิดีโอเข้ามา	55
4.29	โปรแกรมเมื่อทำการแตกไฟล์วิดีโอเป็นเฟรมเสร็จแล้ว	56
4.30	โปรแกรมเมื่อทำการเลือกเฟรม	56
4.31	โปรแกรมเมื่อทำการกดปุ่ม Segmentation	57
4.32	โปรแกรมเมื่อดำเนินการเสร็จสิ้น	57
4.33	โปรแกรมแสดงวิดีโอเมื่อคลิกเลือก news_nbt_01_part02	58

บทที่ 1
บทนำ

1.1 ที่มาและความสำคัญ

ปัจจุบันสื่อมัลติมีเดียได้มีบทบาทต่อชีวิตประจำวันมากขึ้น อันเนื่องมาจากความก้าวหน้าทางด้านเทคโนโลยีคอมพิวเตอร์ซึ่งเกื้อหนุนให้การใช้สื่อมัลติมีเดียเป็นไปอย่างแพร่หลาย ซึ่งสื่อมัลติมีเดียนี้ มีหลายรูปแบบ ไม่ว่าจะเป็นข้อความ , รูปภาพ , เสียง , ภาพนิ่ง , ภาพเคลื่อนไหว , ภาพวิดีโอ เป็นต้น แต่สื่อที่อาจจะเรียกได้ว่าได้รับความนิยมและให้ผลลัพธ์ที่ครบครันต่อการสื่อสารมากที่สุด คงจะได้แก่ ภาพวิดีโอ เนื่องจากเป็นสื่อที่มีทั้งเสียงและภาพ ช่วยให้ผู้ชมสามารถรับสารที่ผู้สื่อสารต้องการจะสื่อได้ตรงประเด็น โดยเฉพาะในโลกไซเบอร์สเปซนั้นมีการนำเข้าและรับชมกันอย่างแพร่หลาย ทำให้นอกเหนือไปจากการใช้ภาพวิดีโอในกิจกรรมนันทนาการส่วนตัวแล้ว ภาพวิดีโอได้ถูกนำไปใช้ในหลายแวดวงที่สำคัญ เช่น ดนตรี , กีฬา , ภาพยนตร์ , วิทยุและโทรทัศน์ , งานด้านสื่อสารมวลชน และงานอื่น ๆ ที่เอื้อต่อธุรกิจขององค์กร

สำหรับภาพวิดีโอ หรือ Digital Video นั้น คือภาพที่เกิดจากการถ่ายด้วย กล้องวิดีโอ , กล้องดิจิตอล , กล้องถ่ายภาพยนตร์ หรืออุปกรณ์ต่าง ๆ ที่เกี่ยวข้อง แล้วนำมาบันทึกให้อยู่ในรูปของไฟล์ โดยใช้ฮาร์ดแวร์ที่เรียกว่า Video Capture Board หรือแผงวงจรจัดการเกี่ยวกับภาพเคลื่อนไหว ในการจับภาพมาเป็นไฟล์ ปัจจุบันไฟล์ภาพวิดีโอมีอยู่หลายชนิดไม่ว่าจะเป็น WMV , AVI , MOV, MPEG ซึ่งก็มีลักษณะคล้ายกันๆ ต่างกันที่ คุณภาพของภาพ , ความต่อเนื่องของภาพ (Playback Rate) , จุดภาพ (Pixel) และขนาดของไฟล์ (Compression) ที่จะมีขนาดเล็กหรือใหญ่แตกต่างกันไป โดยภาพวิดีโอนั้นมีส่วนประกอบที่เรียกว่า เฟรม (Frame) ซึ่งเป็นชุดของภาพนิ่งหลาย ๆ ภาพต่อกันไปตามเวลาของสัญญาณวิดีโอ ในแต่ละเฟรมนั้นอาจแตกต่างกันในส่วนของ ความละเอียดในการแสดงผล (Resolution) , จำนวนสี (Color) , ค่าความแตกต่างระหว่างสีขาวและสีดำ (Contrast) ที่แสดงในแต่ละเฟรม

โครงงานนี้เน้นความสำคัญกับวิดีโอข่าว ซึ่งพบว่าในปัจจุบันผู้ชมนิยมรับข่าวสารผ่านโลกไซเบอร์สเปซมากขึ้น เว็บไซต์ของสำนักข่าวต่าง ๆ หรือสถานีวิทยุโทรทัศน์มีการรายงานข่าวและนำเสนอเป็นภาพวิดีโอผ่านอินเตอร์เน็ตอย่างทันเหตุการณ์ การรายงานข่าวในปัจจุบันเป็นการรายงานข่าวแบบต่อเนื่องไปตลอดรายการ อาจจะมีการเว้นช่วงในการรับชมโฆษณาบ้าง แต่อย่างไรก็ตามก็เป็นการยากในการติดตามชมเฉพาะข่าวที่สนใจในภายหลัง นั่นคือเมื่อผู้ชมต้องการดาวน์โหลดวิดีโอข่าวมาดูย้อนหลัง ก็จะได้ไฟล์วิดีโอที่มีความยาวมากเนื่องจากการรายงานข่าวแต่ละข่าวเป็นไปอย่างต่อเนื่อง ทำให้ไม่สะดวกในการเลือกชม หรือเลือกเก็บไว้ดูเฉพาะข่าวที่สนใจ

ดังนั้นการทำให้ผู้ชมมีความสะดวกในการเลือกรับข่าวสารโดยการทำการแบ่งข่าวตามเนื้อหาของข่าวหนึ่ง ๆ แล้วทำการจัดหมวดหมู่นั้นก็จะเป็นการสร้างความนิยมแก่ผู้ชม และมีผลดีอื่น ๆ ที่ตามมาภายหลัง

โครงงานนี้เป็นการศึกษาค้นคว้าเกี่ยวกับอัลกอริทึมที่จะช่วยในการตัดแบ่งวิดีโอข่าวแบบอัตโนมัติ โดยอาศัยค่าความแตกต่างระหว่างเฟรมของวิดีโอข่าว คือใช้หลักการจับคู่เปรียบเทียบ (Matching) ของ Color layout ได้แก่ ความแตกต่างของระดับสีในแต่ละเฟรม ซึ่งอาจจะมีความแตกต่างมากน้อยต่างกันไป ทำการเปรียบเทียบความแตกต่างของระดับสีในระยะห่างระหว่างเฟรมที่ต่าง ๆ กันไป ซึ่งจะช่วยในการจำแนกวิดีโอข่าวในแต่ละข่าวได้ ทั้งนี้ คุณภาพของภาพวิดีโอทั้งในด้านความคมชัดและจำนวนสีจะมีส่วนช่วยในเรื่องของความแม่นยำในการตัดแบ่งวิดีโอข่าวเป็นอย่างยิ่ง

1.2 วัตถุประสงค์ของโครงการ

1.2.1 เพื่อทำการศึกษาค้นคว้าข้อมูลเกี่ยวกับหลักการประมวลผลและวิเคราะห์ภาพประเภทภาพวิดีโอ

1.2.2 เพื่อออกแบบอัลกอริทึมที่จะใช้สำหรับการวิเคราะห์และตัดแบ่งวิดีโอข่าว ที่มีความยาวหลายข่าวให้มีขนาดที่เล็กลงและมีเนื้อหาเฉพาะ

1.2.3 เพื่อนำอัลกอริทึมที่ออกแบบและพัฒนาได้มาใช้ในการทดลองตัดแบ่งวิดีโอข่าว

1.2.4 เพื่อนำผลที่ได้จากการทดลองมาวิเคราะห์ความถูกต้อง และความแม่นยำในการตัดแบ่งข่าวแต่ละข่าว

1.3 ขอบข่ายของโครงงาน

1.3.1 พัฒนาอัลกอริทึมที่สามารถตัดแบ่งวิดีโอข่าว ให้มีขนาดเล็กลงและมีเนื้อหาเฉพาะเพื่อให้ง่ายในการดูข้อมูล

1.3.2 วิดีโอข่าวที่จะใช้สำหรับโครงงานนี้จะเป็นวิดีโอข่าว ซึ่งตำแหน่งของกล้องและผู้ประกาศข่าวจะต้องมีการเปลี่ยนแปลงไม่มากนัก

1.3.3 ไฟล์วิดีโอข่าวเป็นไฟล์ .wmv ซึ่งเป็นการบีบอัดวิดีโอที่พัฒนาโดยไมโครซอฟท์

1.4 ขั้นตอนการดำเนินงาน

1.4.1 ทำการศึกษาค้นคว้าข้อมูลเกี่ยวกับหลักการประมวลผลและวิเคราะห์ภาพวิดีโอ

1.4.2 ออกแบบอัลกอริทึมที่จะใช้สำหรับการวิเคราะห์องค์ประกอบของภาพและตัดแบ่งวิดีโอข่าวได้เองแบบอัตโนมัติ

1.4.3 ทดลองนำอัลกอริทึมที่ออกแบบและพัฒนาได้ มาใช้ในการตัดแบ่งวิดีโอข่าว

1.4.4 วิเคราะห์ความถูกต้องและค่าความผิดพลาดในการตัดแบ่งข่าวแบบอัตโนมัติ จากผลการทดลองที่ได้

1.4.5 สรุปผลการทดลองและจัดทำรูปเล่มโครงงาน

1.5 แผนการดำเนินงาน

ตารางที่ 1.1 แผนการดำเนินงาน

กิจกรรม	พ.ศ. 2551						พ.ศ. 2552		
	ก.ค.	ส.ค.	ก.ย.	ต.ค.	พ.ย.	ธ.ค.	ม.ค.	ก.พ.	มี.ค.
1. ทำการศึกษาข้อมูลเกี่ยวกับทฤษฎีการประมวลผลและวิเคราะห์ภาพ									
2. ออกแบบอัลกอริทึมที่จะใช้สำหรับการวิเคราะห์และตัดแบ่งวิดีโอข่าว									
3. ทดลองนำอัลกอริทึมที่พัฒนาได้มาใช้ในการตัดแบ่งวิดีโอข่าว									
4. นำผลที่ได้จากการทดลองมาวิเคราะห์ความถูกต้องและค่าความผิดพลาดในการตัดแบ่งข่าว									
5. สรุปผลการทดลองและจัดทำรูปเล่มโครงงาน									

1.6 รายละเอียดงบประมาณของโครงงาน

1.6.1 ค่าใช้จ่ายระหว่างจัดทำโครงงาน 1,000 บาท

1.6.2 ค่าใช้จ่ายในการจัดทำรูปเล่มรายงาน 1,400 บาท

รวมเป็นเงินทั้งสิ้น 2,400 บาท

(สองพันสี่ร้อยบาทถ้วน)

หมายเหตุ ขออนุมัติถัวเฉลี่ยทุกรายการ

บทที่ 2
ทฤษฎีที่เกี่ยวข้อง

2.1 การตัดแบ่งวิดีโอ (Video Segmentation)

2.1.1 ความรู้พื้นฐานเกี่ยวกับวิดีโอ

ปัจจุบันมัลติมีเดียจัดว่าเป็นสื่ออีกชนิดหนึ่งที่ได้รับความนิยมใช้งานกันอย่างแพร่หลาย ไม่ว่าจะเป็นการใช้เป็นสื่อบันเทิง (Entertainment) การนำเสนอผลิตภัณฑ์และบริการต่างๆ (Product and Service Presentation) การเรียนการสอนผ่านระบบอิเล็กทรอนิกส์ (E-learning) และการนำเสนอผลงานต่าง ๆ (Task Presentation) จึงนับว่าเป็นสื่อที่มีอิทธิพลต่อสังคมเป็นอย่างมาก

มัลติมีเดีย คือ ระบบสื่อสารข้อมูลข่าวสารหลายชนิด โดยผ่านสื่อทางคอมพิวเตอร์ซึ่งประกอบด้วย ข้อความ ฐานข้อมูล ตัวเลข กราฟิก ภาพเสียง และภาพวิดีโอ

ภาพวิดีโอเป็นองค์ประกอบของมัลติมีเดียที่มีความสำคัญเป็นอย่างมาก เนื่องจากวิดีโอในระบบดิจิตอลสามารถนำเสนอข้อความหรือรูปภาพ (ภาพนิ่งหรือภาพเคลื่อนไหว) ประกอบกับเสียงได้สมบูรณ์มากกว่าองค์ประกอบชนิดอื่นๆ อย่างไรก็ตาม ปัญหาหลักของการใช้วิดีโอในระบบมัลติมีเดียก็คือ การสิ้นเปลืองทรัพยากรของพื้นที่บนหน่วยความจำเป็นจำนวนมาก เนื่องจากการนำเสนอวิดีโอด้วยเวลาที่เกิดขึ้นจริง (Real-Time) จะต้องประกอบด้วยจำนวนภาพไม่ต่ำกว่า 30 ภาพต่อวินาที (Frame/Second) ถ้าหากการประมวลผลภาพดังกล่าวไม่ได้ผ่านกระบวนการบีบอัดขนาดของสัญญาณมาก่อน การนำเสนอภาพเพียง 1 นาทีอาจต้องใช้หน่วยความจำมากกว่า 100 เม็กกะไบต์ (MB) ซึ่งจะทำให้ไฟล์มีขนาดใหญ่เกินขนาดและมีประสิทธิภาพในการทำงานที่ด้อยลง ซึ่งเมื่อมีการพัฒนาเทคโนโลยีที่สามารถบีบอัดขนาดของภาพอย่างต่อเนื่องจนทำให้ภาพวิดีโอสามารถทำงานได้อย่างมีประสิทธิภาพมากยิ่งขึ้นและกลายเป็นสื่อที่มีบทบาทสำคัญต่อระบบมัลติมีเดีย (Multimedia System) [1]

วิดีโอ (Video) แบ่งเป็น 2 ชนิด คือ วิดีโออนาลอก (Analog Video) และวิดีโอดิจิตอล (Digital Video) วิดีโออนาลอกเป็นวิดีโอที่ทำการบันทึกข้อมูลภาพและเสียงให้อยู่ในรูปของสัญญาณอนาลอก (ในรูปของคลื่น) ได้แก่ VHS (Video Home System) ส่วนวิดีโอดิจิตอลเป็นวิดีโอที่ทำการบันทึกข้อมูลภาพและเสียงที่ได้มาจากกล้องวิดีโอดิจิตอล ให้อยู่ในรูปสัญญาณดิจิตอลคือ 0 กับ 1 ซึ่งทำให้สามารถบันทึกลงบนฮาร์ดดิสก์ ซีดีรอม ดีวีดี หรืออุปกรณ์บันทึกข้อมูลอื่น ๆ ได้

การพัฒนาของวิดีโอดิจิตอลส่งผลให้วิดีโออนาลอกหายไปจากวงการมัลติมีเดีย เนื่องจากสัญญาณดิจิตอล สามารถที่จะบันทึกข้อมูลลงบนฮาร์ดดิสก์ ซีดีรอม ดีวีดี หรืออุปกรณ์บันทึกข้อมูล

อื่นๆ และสามารถแสดงผลบนคอมพิวเตอร์ได้อย่างมีประสิทธิภาพ ในการผลิตมัลติมีเดียบนคอมพิวเตอร์ สามารถเปลี่ยนรูปแบบของสัญญาณอนาลอกเป็นสัญญาณดิจิตอลได้ เพียงแต่ผู้ผลิตมีทรัพยากรทางด้านฮาร์ดแวร์และซอฟต์แวร์ที่เหมาะสมเท่านั้น

2.1.1.1 คุณภาพของวิดีโอ

การวัดคุณภาพของวิดีโอสามารถวัดได้จาก อัตราเฟรม (Frame Rate) และความละเอียด (Resolution) ของภาพ

1. อัตราเฟรม (Frame Rate) คืออัตราความถี่ในการแสดงภาพจากไทม์ไลน์ (Timeline) ออกจากหน้าจอ อัตราที่เฟรมถูกแสดงในวิดีโอมีหน่วยเป็นเฟรมต่อวินาที หรือ fps (Frame Per Second) เป็นหน่วยวัดปริมาณข้อมูลที่ ใช้ในการเก็บบันทึกและแสดงวิดีโอ นอกจากนี้ยังสามารถแสดงเฟรมให้มีความต่อเนื่องในเวลาอันรวดเร็วโดยผู้จัดทำสามารถที่จะกำหนดอัตราเฟรมเองได้ เช่น อัตราเฟรมของภาพยนตร์เท่ากับ 24 เฟรมต่อวินาที อัตราเฟรมโทรทัศน์ระบบ PAL เท่ากับ 25 เฟรมต่อวินาที และอัตราเฟรมโทรทัศน์ระบบ NTSC เท่ากับ 30 เฟรมต่อวินาที

2. *ความละเอียด (Resolution)* หมายถึง ความคมชัดของภาพที่แสดงผลออกทางจอภาพ ความละเอียดของจอภาพขึ้นอยู่กับจำนวนจุดทั้งหมดที่เกิดบนจอ จุดต่างๆ นี้ เรียกว่า พิกเซล (Pixels) เช่น ความละเอียดของจอภาพ 640 X 480 พิกเซลหมายความว่า มีจำนวนจุด แสดงผลเรียงกันอยู่บนจอภาพในแนวนอน 640 พิกเซล และแนวตั้ง 480 พิกเซล ซึ่งรวมพิกเซลทั้งหมดเป็น 370,200 พิกเซล นอกจากนี้ภาพที่มีขนาดเท่ากัน บางครั้งก็อาจจะมีความละเอียดที่ต่างกันได้ เนื่องจากจำนวนพิกเซลต่างกันจะส่งผลให้ขนาดของพิกเซลต่างกันด้วย ยิ่งความละเอียดของจอภาพสูงจะยิ่งทำให้มองเห็นพื้นที่ใช้งานบนจอภาพกว้างมากขึ้น แต่จะทำให้มีขนาดเล็กลง

2.1.1.2 มาตรฐานของภาพวิดีโอ

1. NTSC (National Television System Committee) เป็นการเข้ารหัสข้อมูลแบบสัญญาณอิเล็กทรอนิกส์ ที่กำหนดให้แสดงภาพด้วยเส้นในแนวนอน 525 เส้นต่อเฟรม ในอัตรา 30 เฟรมต่อวินาที มีสี 16 ล้านสี ขณะที่จอคอมพิวเตอร์จะใช้วิธีการที่เรียกว่า Progressive-Scan ใช้ในสหรัฐอเมริกา และญี่ปุ่น

2. PAL (Phase Alternate Line) เป็นการสร้างภาพจากเส้นแนวนอน 625 เส้นต่อเฟรม ด้วยอัตรา 25 เฟรมต่อวินาที และทำการแสดงภาพด้วยวิธี Interlacing ใช้ในแถบยุโรป อังกฤษ ออสเตรเลีย แอฟริกาใต้ และประเทศไทย

3. SECAM (Sequential Color and Memory) จะแพร่สัญญาณแบบอนาลอก ส่วนการสร้างภาพเป็น 819 เส้น ด้วยอัตราการรีเฟรช 25 เฟรมต่อวินาทีใช้ในประเทศฝรั่งเศส รัสเซีย ยุโรปตะวันออกและตะวันออกกลาง

4. HDTV (High Definition Television) เป็นเทคโนโลยีของการแพร่ภาพที่ถูกพัฒนาเพื่อแสดงภาพที่มีความละเอียดสูง คือ 1280x720 พิกเซล เป็นความละเอียดเช่นเดียวกับภาพในโรงภาพยนตร์ ซึ่งถูกพัฒนาออกมาใช้ครั้งแรกในปี 1998

2.1.1.3 การบีบอัดวิดีโอ

1. JPEG เป็นการย่อภาพสีให้คงรายละเอียดเดิมไว้ให้มากที่สุด ซึ่งเป็นการยุบพื้นที่ขนาด 8x8 พิกเซลให้เหลือเป็น 1 พิกเซล โดยเลือกสีที่มีอยู่มากที่สุดเป็นสีสำหรับพิกเซลนั้น อัตราการบีบอัด 25:1 40:1 จนถึง 100:1

2. Windows Media Video (WMV) เป็นรูปแบบการบีบอัดวิดีโอที่พัฒนาโดยไมโครซอฟท์ ซึ่งออกแบบมาสำหรับแอพพลิเคชันบนอินเตอร์เน็ตสตรีมมิ่ง

3. CODEC เป็นการบีบอัดที่สามารถนำไปใช้กับ ซอฟต์แวร์และฮาร์ดแวร์ ส่วนใหญ่นิยมใช้ในการบีบอัดแบบ MPEG , Indeo และ Cinepak

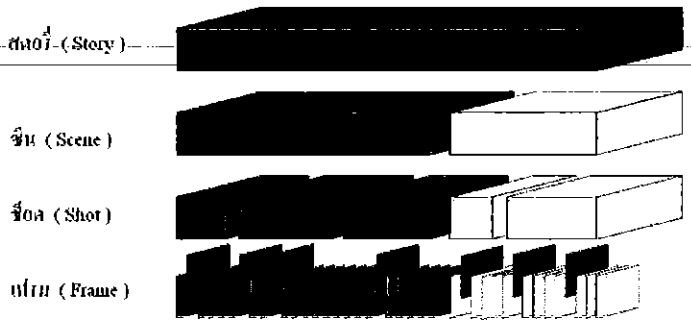
4. MPEG (Moving Picture Experts Group) เป็นการบีบอัดที่คล้ายกับแบบ JPEG แต่จะลดจำนวนข้อมูลที่ซ้ำกันของภาพต่อไป เป็นการบีบอัดแบบไม่สมมาตร คือขั้นตอนในการเข้ารหัสสัญญาณวิดีโอจะนานกว่าขั้นตอนการถอดรหัสข้อมูล

5. DivX พัฒนาจากระบบ CODEC สามารถลดข้อมูล เหลือเพียง 10-20% ของปริมาณข้อมูลเดิม [2]

2.1.2 ความรู้พื้นฐานเกี่ยวกับการตัดแบ่งวิดีโอ (Video Segmentation)

2.1.2.1 องค์ประกอบของภาพวิดีโอ

โดยปกติแล้วภาพวิดีโอที่มีความยาวต่อเนื่องกันนาน ๆ จะทำให้เกิดความลำบากในการเลือกชมในช่วงที่ต้องการ การตัดภาพวิดีโอออกเป็นส่วน ๆ ตามเนื้อหาจะช่วยให้ไฟล์ภาพวิดีโอมีความกะทัดรัดและแบ่งประเภทได้ง่ายขึ้น อาจกล่าวได้ว่าภาพวิดีโอ แบ่งออกเป็น 4 ลำดับชั้น คือ



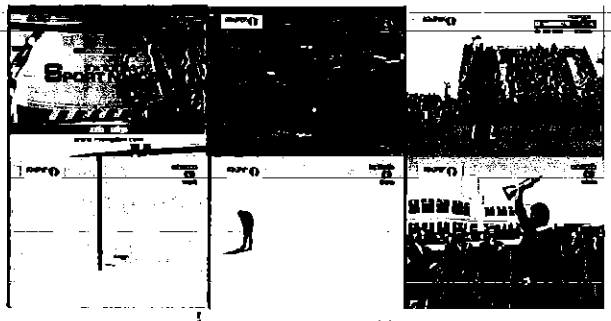
รูปที่ 2.1 องค์ประกอบของภาพวิดีโอ [3]

1. สตอรี่ (Story) เป็นส่วนของภาพวิดีโอที่ใหญ่ที่สุด หมายถึง ตอนหรือช่วงเหตุการณ์หนึ่งเป็นการรวบรวมเอาซีนหลาย ๆ ซีนที่มีความสัมพันธ์กันมาต่อเนื่องกันเข้า และเมื่อรวมต่อเข้ากันแล้วจะเกิดผลสมบูรณ์ของเนื้อหาอยู่ในตัวเอง สามารถจบ เหตุการณ์ในช่วงนั้น ๆ โดยที่ผู้ชมเข้าใจได้ ซึ่งในสตอรี่หนึ่งๆ อาจประกอบด้วยซีนเดียวหรือหลายซีนได้

รูปที่ 2.2 สตอรี่ของวิดีโอข่าว

2. ซีน (Scene) เป็นส่วนประกอบย่อยของสตอรี่อีกทีหนึ่ง เป็นการนำเอาช็อตหลาย ๆ ช็อตมารวมกัน ซึ่งเป็นช็อตที่เหตุการณ์เกิดขึ้นในสถานที่ เดียวกัน เวลาเดียวกัน หรือมีความต่อเนื่องทางเนื้อหา นำเอาช็อตต่าง ๆ มาเรียงกันเป็นเรื่องราว เพื่อให้การใช้ภาษาภาพ สามารถถ่ายทอดความคิดเนื้อหา เป็นไปอย่างมีระบบ มีลำดับที่ทำให้ผู้ชมเข้าในแจ่มแจ้ง

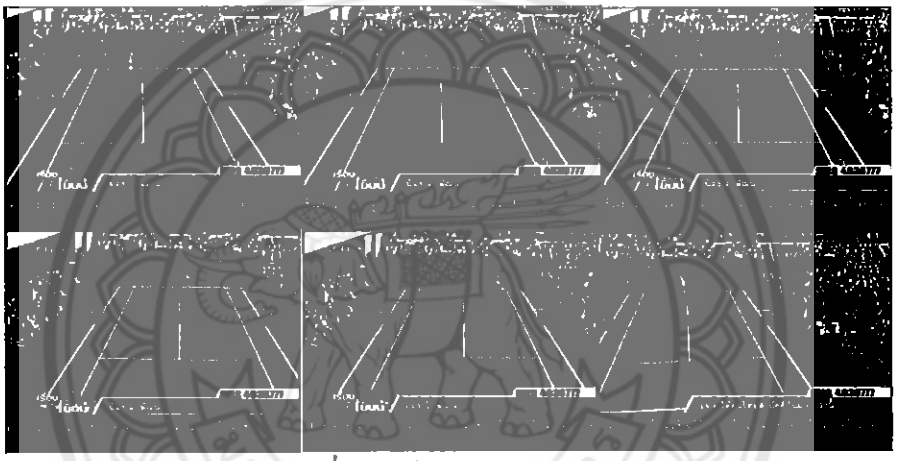
ในแต่ละซีน อาจมีหลายช็อตหรือช็อตเดียวได้ ขึ้นอยู่กับความเข้าใจของผู้ชม กรณีที่ซีนเกิดจากช็อตเพียงช็อตเดียวนั้น อาจจะเป็นเหตุการณ์สั้น ๆ ไม่เกิน 1 นาที แต่โดยทั่วไป ซีนของภาพวิดีโอมักเกิดจากหลาย ๆ ช็อตมารวมกันเป็นส่วนมาก ซึ่งช็อตอาจจะเกิดจากกล้องวิดีโอที่บันทึกภาพวิดีโอมากกว่า 1 ตัวขึ้นไป



รูปที่ 2.3 ซีนของวิดีโอข่าว

3. ช็อต (Shot) หมายถึง ลักษณะภาพที่เกิดจากการถ่ายภาพวิดีโอตั้งแต่เริ่มถ่ายไปจนถึงการหยุดการเดินกล้อง เรียกว่า 1 ช็อต นั่นคือแต่ละช็อตของภาพวิดีโอจะมีความต่อเนื่องกัน หรืออาจจะกล่าวได้ว่าเป็นช่วงที่ภาพวิดีโอเกิดจากการถ่ายจากกล้องวิดีโอตัวใดตัวหนึ่งอย่างต่อเนื่อง ซึ่งแต่ละช็อตจะเกิดจากการบันทึกภาพของกล้องวิดีโอเพียงตัวเดียว การเปลี่ยนจากช็อตหนึ่งไปช็อตหนึ่งของภาพวิดีโอนี้ เกิดขึ้นเร็วมากแบบทันทีทันใดในชั่วระยะที่กะพริบตาจนบางครั้งเราอาจไม่ทันสังเกต เรียกว่าภาพวิดีโอจะรันยาวต่อเนื่องโดยไม่มีการขัดจังหวะ (Uninterrupted) ทั้งนี้เพื่อความต่อเนื่องของผู้รับชมนั่นเอง

ความยาวของแต่ละช็อต อาจจะเป็นแค่ 2 วินาทีไปจนถึง 30 วินาที ไม่มีกฎตายตัวแน่นอนเกี่ยวกับความยาวในแต่ละช็อตทั้งนี้ขึ้นอยู่กับรายละเอียดของภาพวิดีโอนั้น ๆ



รูปที่ 2.4 ช็อตของวิดีโอข่าว

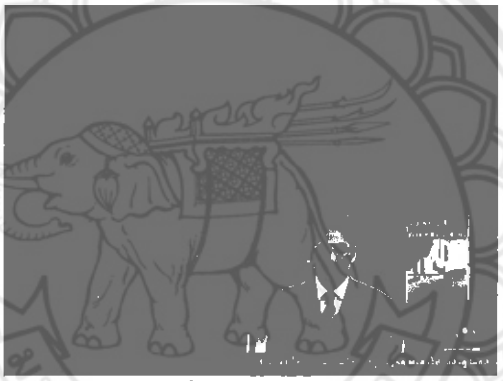
ในทางเทคนิคมีนิยามของคำว่า ช็อต คือ "Created of a series of frames" นั่นคือการรวมกันของเฟรมย่อย ๆ แต่ละเฟรมจะเกิดขึ้นเป็นช็อตหนึ่ง ๆ นั่นเอง

เทคนิคการเชื่อมต่อช็อตต่าง ๆ ให้เป็นซีน ทำได้หลายวิธี ได้แก่

3.1 คัต (Cut) หรือ การตัดแบบทันทีทันใด (Straight Cut) คือการเชื่อมโยงระหว่างช็อตแบบตัดไปตรง ๆ เหมือนเอาภาพของ 2 ช็อตมาต่อกันเฉย ๆ ซึ่งดูเหมือนว่า จะสะดุดอารมณ์คนดูแต่จริง ๆ แล้ว หากเลือกคัตภาพในจังหวะที่เหมาะสมกับความต้องการทางความหมาย และอารมณ์ที่ต้องการจะสื่อให้ผลดีมาก เพราะที่จริงแล้วคนเราเวลารับรู้เหตุการณ์ต่าง ๆ มิใช่จะเปิดรับเหตุการณ์แบบยาวต่อเนื่องกันตลอด แต่จะเป็นการเลือกมองเฉพาะส่วนที่ต้องการจะรับรู้เท่านั้น ส่วนที่ไม่น่าสนใจ หรือไม่เกี่ยวข้องกับการรับรู้ จะถูกตัดออกไป ทำให้วิธีการเชื่อมโยงระหว่างช็อตแบบคัตทำให้เกิดความรู้สึกเป็นธรรมชาติมากที่สุด และโดยทั่วไปแล้วในภาพยนตร์และโทรทัศน์จะพบได้บ่อยที่สุดกว่าการใช้วิธีอื่น ๆ

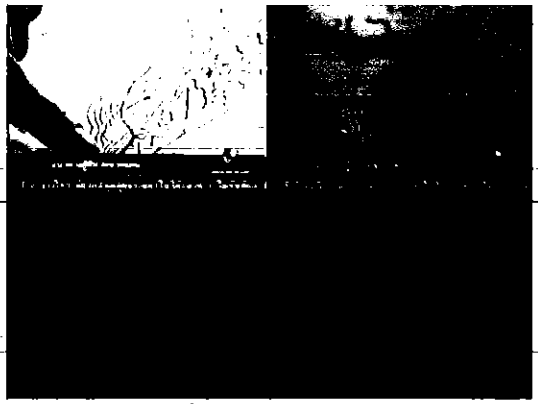
รูปที่ 2.5 เทคนิคการเชื่อมต่อช็อตแบบคัต (Cut)

3.2 เฟด (Fade) มี 2 อย่าง คือ เฟดอิน (Fade In) หมายถึง การเริ่มต้นภาพจากเฟรมที่มืดสนิท หรือพื้นสี (Color Background) แล้วค่อย ๆ ปรากฏภาพให้เห็นชัดเจนขึ้นจนเป็นปกติ ส่วนใหญ่จะนำมาใช้บอกการเริ่มต้นของเรื่อง ของเหตุการณ์ หรือของวันใหม่ เป็นต้น



รูปที่ 2.6 เทคนิคการเชื่อมต่อช็อตแบบเฟดอิน (Fade In)

ส่วนเฟดเอาต์ (Fade Out) หมายถึง การนำภาพที่กำลังมองเห็นอยู่ชัดเจนให้ค่อย ๆ จางหายสู่ความมืด หรือพื้นสี ในที่สุด ส่วนใหญ่ใช้เพื่อการบอกการจบสิ้นเรื่องของเหตุการณ์ หรือสิ้นสุดลง เป็นต้น การใช้เฟดนี้ จะให้อารมณ์แบบค่อยเป็นค่อยไปอย่างนุ่มนวลกว่าคัต บางครั้งเฟดอิน - เฟดเอาต์ ถูกใช้เพื่อการเปลี่ยนฉาก เหตุการณ์ หรืออารมณ์ อย่างรวดเร็ว เพื่อให้เรื่องราวดำเนินไปได้โดยที่คนดูจะไม่รู้สึกว่ามีการสะดุด หรือที่เรียกว่า ภาพกระโดด (Jumping cut)



รูปที่ 2.7 เทคนิคการเชื่อมต่อช็อตแบบเฟดเอาต์ (Fade Out)

3.3 ดิสโซลว์ (Dissolve) คือ การที่ภาพใน ช็อต หนึ่งที่กำลังจางหายไป ก็มีภาพในอีก ช็อต มาซ้อนแล้วค่อย ๆ ชัดขึ้น และมาแทนที่ในที่สุด การใช้ดิสโซลว์นี้ใช้เพื่อเชื่อมโยงเหตุการณ์ระหว่างช็อต แล้วมีความสัมพันธ์ที่เกี่ยวเนื่องกันอย่างกลมกลืน หรือเอาใช้เพื่อลัดเวลา ซึ่งจะทำให้ความรู้สึกขณะรับชมเป็นแบบนุ่มนวล ชวนฝัน

รูปที่ 2.8 เทคนิคการเชื่อมต่อช็อตแบบดิสโซลว์ (Dissolve)

3.4 ไวป์ (Wipe) คือการกวาดภาพ นำเอาภาพใหม่แทนที่ภาพเก่า เหมือนกับการเปิดปิดม่านเวทีละคร นำมาใช้เพื่อเล่าเรื่องราวแบบตรงไปตรงมา ไม่ต้องการความสมจริง สอดคล้องกับธรรมชาติการรับรู้ของมนุษย์เรา

รูปที่ 2.9 เทคนิคการเชื่อมต่อช็อตแบบไวป์ (Wipe)

ซึ่งเทคนิคต่าง ๆ ทั้ง 4 แบบนี้ อาจมีผลต่อการตัดต่อภาพวิดีโอแบบอัตโนมัติ โดยเฉพาะอย่างยิ่งการเชื่อมต่อช็อตต่าง ๆ ให้เป็นซีน แบบเฟดและดิสโซลว์จะมีการเปลี่ยนช็อตอย่างนุ่มนวล ทำให้องค์ประกอบของภาพมีความคล้ายคลึงกัน [4]

4. เฟรม (Frame) เป็นส่วนย่อยที่สุดของภาพวิดีโอ แต่ละเฟรมย่อย ๆ จะประกอบกันเป็นช็อตหนึ่ง ๆ จะมีคีย์เฟรม (Key Frame) คือ เฟรมที่มีวัตถุหรือมีการเปลี่ยนแปลง สามารถสังเกตได้ ซึ่งจะใช้ในการเปรียบเทียบความแตกต่างระหว่างเฟรมได้

2.2 ระบบสี Y/Cb/Cr [5]

Y/Cb/Cr เป็นระบบสีพื้นฐานที่นิยมใช้กันในภาพดิจิตอลหรือระบบวิดีโอ ระบบสีนี้ถูกพัฒนาขึ้นโดยสหภาพโทรคมนาคม หรือ ITU (International Telecommunication Union) เพื่อใช้เป็นมาตรฐานสำหรับวิดีโอในระบบดิจิตอลทั่วโลก

ระบบ Y/Cb/Cr จะแทนสีด้วยความสว่าง (Y) และสองช่องสัญญาณสี (Cb/Cr) โดย ค่า Y แทนความสว่าง (Luminance) และค่า Cb และ Cr จะเป็นค่าความต่างเทียบกับ Y ของสีน้ำเงิน (Blue) และสีแดง (Red) ตามลำดับ กล่าวคือ Cb คือสีน้ำเงินที่ตัด Luminance ออกไป Cr คือสีแดงที่ตัด Luminance ออกไป

2.3 MPEG-7 [6]

MPEG-7 เป็นมาตรฐาน ISO/IEC (International Organization for Standard / International Electrotechnical Commission Standard) ซึ่งถูกพัฒนาโดย Moving Pictures Experts Group (MPEG) ซึ่งเป็นกลุ่มที่ทำงานใน ISO มาตรฐาน MPEGดังกล่าวร่วมกันได้รวมถึงสามารถแลกเปลี่ยนข้อมูลดังกล่าวผ่านเครือข่ายชนิดต่างๆ ได้ MPEG-7 ถูกกำหนดขึ้นในปี 2001 มีชื่ออย่างเป็นทางการคือ Multimedia Content Description Interface มีจุดมุ่งหมายเพื่ออธิบายเนื้อหา (Content) ของข้อมูลมัลติมีเดีย (Multimedia Data) โดยการผูก Metadata เข้ากับเนื้อหาของข้อมูล

มัลติมีเดีย MPEG-7 กำหนดมาตรฐานชุดของเครื่องมือซึ่งใช้ในการอธิบายสารสนเทศมัลติมีเดียไว้หลายชนิด MPEG-7 ไม่ได้ถูกพัฒนามาเพื่อแทนที่มาตรฐานเดิมของ MPEG แต่มีจุดประสงค์เพื่อเตรียมฟังก์ชันการทำงานเพิ่มเติมให้กับ MPEG มาตรฐานอื่น เพื่อกำหนดรูปแบบและวิธีการในการอธิบายเนื้อหาของข้อมูลมัลติมีเดียให้มีรูปแบบและมาตรฐานเดียวกัน เพื่อให้ผู้ใช้สามารถใช้ข้อมูลสะดวกขึ้น มาตรฐาน MPEG-7 ช่วยให้ผู้ใช้สามารถที่จะทำการ Browsing , Editing , Searching และอื่นๆ รวมถึงช่วยให้ผู้ใช้ที่เป็นเจ้าของข้อมูลสามารถจัดการกับข้อมูลของตนได้อย่างมีประสิทธิภาพและเกิดประโยชน์มากยิ่งขึ้น เช่น การทำดัชนี (Indexing) , การจัดกลุ่มข้อมูล (Data Clustering) , การจำแนกข้อมูล (Data classification) , การบีบอัดข้อมูล(Compression)

2.3.1 ทฤษฎี Color Layout Descriptor ของ MPEG-7 [7]

MPEG-7 มี Descriptor ที่ใช้อธิบายเนื้อหาของข้อมูลโดยใช้ลักษณะต่างๆ เช่น สี (Color) , รูปร่าง (Shape), พื้นผิว(Texture) เป็นต้น ซึ่งสิ่งต่างๆ เหล่านี้เป็น MPEG-7 Multimedia Description Schemes

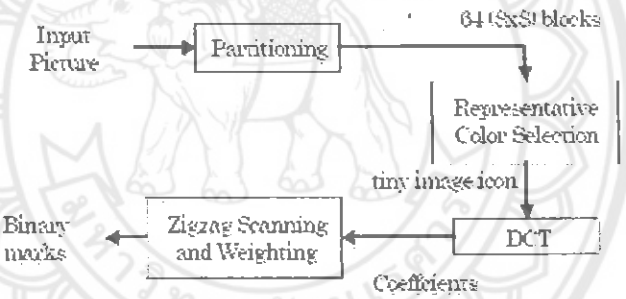
Color Descriptor เป็นการอธิบายเนื้อหาของรูปภาพโดยใช้หลักการของสี ในมาตรฐาน MPEG-7 ประกอบด้วย Scalable Color , Dominant Color , Color Structure และ Color Layout

Color Layout เป็น Descriptor ที่มีความกระชับและช่วยให้การค้นคืนภาพทำได้เร็วขึ้น ออกแบบมาเพื่อเพิ่มประสิทธิภาพในการอธิบายการกระจายตัวของสี (Spatial Distribution of Color) ฟีเจอร์นี้สามารถใช้ได้กับการใช้หลักการเปรียบเทียบความเหมือนในการค้นคืนภาพ (Similarity-based retrieval) , การกรองเนื้อหา (Content Filter) และ Visualization โดยเฉพาะอย่างยิ่งจะมีประโยชน์สำหรับ spatial-structure based retrieval application ตัวอย่างเช่น Sketch based retrieval , แสดงการตัดวิดีโอ (Video Segmentation Identification) วิธีการ Sketch based retrieval จะใช้สำหรับฟังก์ชันที่มีความสำคัญมาก ๆ เนื่องจากมีความสามารถที่ช่วยผู้ใช้อย่างมาก โดยเฉพาะอย่างยิ่งเมื่อการค้นหามีความเร็วเพียงพอ การทำงานของ Descriptor ตัวนี้คือ การจับคู่ (Matching) ระหว่างภาพกับภาพ วิดีโอกับวิดีโอ และ ภาพร่างกับภาพหรือวิดีโอคลิป Description ของ Color Layout ได้มาจากการใช้ข้อมูล Grid Layout ของ MPEG-7 และ Dominant Color Descriptor อย่างไรก็ตามถ้าใช้การรวมกันของทั้งสองตัวนี้จะใช้จำนวนของบิต มาก และการจับคู่ (Matching) จะมีความซับซ้อนมาก จึงสร้าง Color Layout Descriptor เพื่อให้ความถูกต้องและรวดเร็วในการค้นคืนภาพแต่เป็น Descriptor มีความกระชับกว่า

2.3.1.1 การดึงข้อมูลจากภาพในระบบสี Y/Cb/Cr (Extraction)

Descriptor นี้ได้มาโดยการใช้ DCT (Discrete Cosine Transform) บนอาร์เรย์ 2 มิติของการแสดงพื้นที่สีในระบบสี Y/Cb/Cr ดังรูปที่ 2.10 แสดงกระบวนการของ Descriptor ในการดึงข้อมูล

จากภาพ ซึ่งประกอบไปด้วย 4 ขั้นตอนคือ การแบ่งภาพ (Image Partitioning) , การตรวจจับสี (Representative Color Detection) , DCT , และ การประเมินค่าแบบไม่เป็นเชิงเส้น (Non-Linear Quantization) ของ Zigzag-Scanned Coefficient ขั้นแรกภาพจะถูกแบ่งเป็น 64 บล็อก จากนั้นจะถูกเลือกมา 1 ค่าจากแต่ละบล็อก การเลือกการแสดงค่าสีจะใช้วิธีใดก็ได้ แต่จะแนะนำให้ใช้ค่าเฉลี่ยสีของ พิกเซล (Average of pixel color) ในการแสดงค่าสี เนื่องจากเป็นวิธีที่ง่ายที่สุดและมีความถูกต้องเพียงพอ ผลของการเลือกจะให้ภาพเล็กๆที่มีขนาด 8x8 ในขั้นตอนที่ 3 แต่ละองค์ประกอบของสีทั้ง 3 อย่างจะถูกแปลงค่าโดยใช้ DCT 8x8 ดังนั้นจะได้สัมประสิทธิ์ของ DCT 64 ตัวจำนวน 3 ชุด ซึ่งก็คือ Zigzag-Scanned Coefficient และเป็น สัมประสิทธิ์ (Coefficient) ตัวแรกที่มีขนาดเล็ก (สำหรับสัมประสิทธิ์ DC และ AC ใช้ 64 และ 32 ระดับ) มาตรฐานนี้ใช้การอธิบายฟีเจอร์ โดยการควบคุมจำนวนของสัมประสิทธิ์ โดยส่วนใหญ่ใช้สัมประสิทธิ์ 12 ตัว สำหรับ Luminance 6 , Chrominance 3 อย่างไรก็ตามอาจใช้ 18 ตัวได้ (Luminance 6 , Chrominance 6) สำหรับอธิบายภาพนิ่งที่มีคุณภาพสูง จำนวนความยาวของบิตทั้งหมด (Total bit-length) ของ Descriptor (12 Coefficient) คือ 64 บิต รวมบิตสัญญาณ (Signal bit) ซึ่งจะระบุจำนวนของสัมประสิทธิ์



รูปที่ 2.10 กระบวนการของ Descriptor ในการดึงข้อมูลจากภาพ

จำนวนสัมประสิทธิ์ของ DCT ที่ใช้ใน Color Layout Descriptor จะไม่คงที่และจะถูกแสดงใน Coefficient Pattern field ซึ่งสามารถเป็นไปได้ 3 ค่า คือ ค่าแรกใช้ 6 DCT Coefficient สำหรับ Luminance และ สำหรับแต่ละ Chrominance ใช้ 3 DCT Coefficient , ค่าที่สองใช้ 6 DCT Coefficient ทั้ง Luminance และ Chrominance , และสุดท้ายจำนวนสัมประสิทธิ์ของ DCT จะถูกอธิบายใน Number Of YCoefficient และ Number Of CCoefficient Field จำนวนสัมประสิทธิ์ของ DCT ที่เป็นไปได้ คือ 3 , 6 ,10 , 15 ,21 ,28 และ 64 ค่าที่แท้จริงจะถูกแสดงในอาร์เรย์ Ycoeff , CbCoeff และ CrCoeff จะมีความยาว 5 หรือ 6 บิตขึ้นอยู่กับสัมประสิทธิ์

ตารางที่ 2.1 ตารางแสดงรายละเอียดข้อมูลสัมประสิทธิ์

ข้อมูล	จำนวนบิต	คำอธิบาย
CoefficientPattern	1-2	ระบุจำนวนสัมประสิทธิ์ DCT
NumberOfYCoefficient	3	จำนวนสัมประสิทธิ์ DCT ของ luminance
NumberOfCCoefficient	3	จำนวนสัมประสิทธิ์ DCT ของ chrominance
Ycoeff	5-6	ค่าสัมประสิทธิ์ DCT ของ luminance
CbCoeff	5-6	ค่าสัมประสิทธิ์ DCT ของ chrominance
CrCoeff	5-6	ค่าสัมประสิทธิ์ DCT ของ chrominance

2.3.1.2 การจับคู่เปรียบเทียบ (Matching)

Descriptor นี้ใช้ได้กับทั้งหมดของภาพและบางส่วนของภาพจากการระบุรูปร่าง ในการใช้การเจาะจงรูปร่างการเลือกการแสดงค่าสีควรใช้เฉพาะพิกเซลที่ถูกต้องเท่านั้น ซึ่งต้องทำก่อนขั้นตอน DCT transform

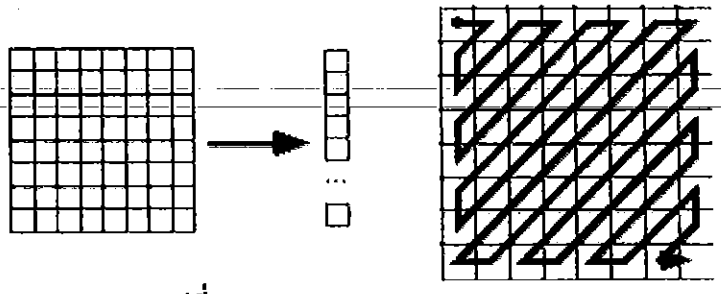
สำหรับการจับคู่เปรียบเทียบ (Matching) ของ Color Layout ระหว่าง {DY , DCr , DCb} และ {DY , DCr' , DCb'} ใช้สมการต่อไปนี้ในการหาค่าความแตกต่าง (Distance)

$$D = \sqrt{\sum_i w_{yi}(DY_i - DY_i')^2} + \sqrt{\sum_i w_{bi}(DCb_i - DCb_i')^2} + \sqrt{\sum_i w_{ri}(DCr_i - DCr_i')^2} \quad (2.1)$$

ซึ่ง i คือลำดับของสัมประสิทธิ์ Zigzag-Scanned ระบบลักษณะการมองเห็นของมนุษย์สามารถคำนวณได้เหมือนกันเนื่องจากฟีเจอร์นี้จะอธิบายในเชิงความถี่ (frequency domain)

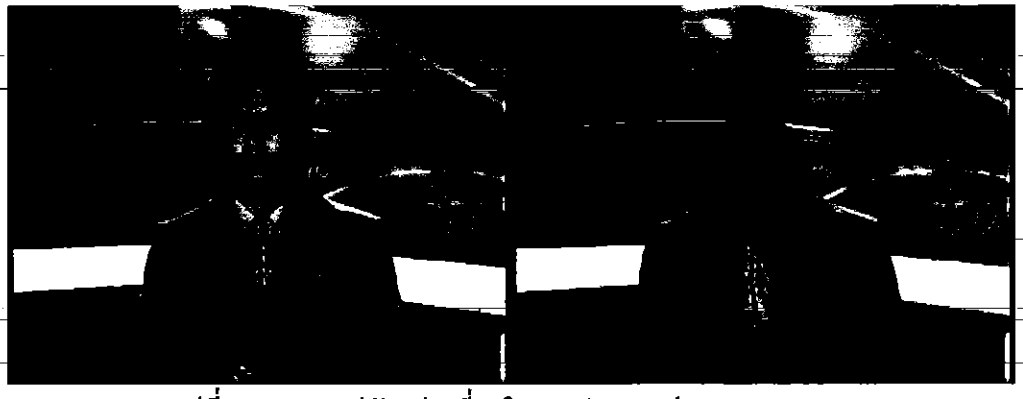
จากสมการการจับคู่เปรียบเทียบ (Matching) ของ Color Layout สมการที่ 2.1

ใช้การสแกนแบบ Zig-zag มีลักษณะดังนี้



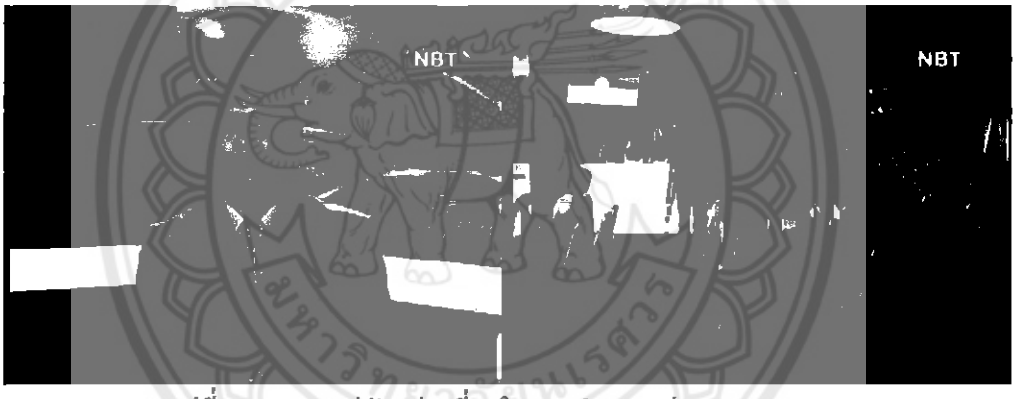
รูปที่ 2.11 การสแกนแบบ Zig-zag

2.3.1.3 ตัวอย่างการคำนวณ Color Layout ของภาพตัวอย่าง



รูปที่ 2.12 ภาพคู่ตัวอย่างที่ 1 ในการคำนวณค่า Color Layout

จากรูปที่ 2.12 จะได้ค่า Color Layout จากการคำนวณคือ 5.5373191



รูปที่ 2.13 ภาพคู่ตัวอย่างที่ 2 ในการคำนวณค่า Color Layout

จากรูปที่ 2.13 จะได้ค่า Color Layout จากการคำนวณคือ 78.399345

จากผลลัพธ์ของคู่ตัวอย่างทั้งสองคู่ พบว่า ภาพที่มีลักษณะใกล้เคียงกัน จะมีค่าความแตกต่าง (Distance) ต่อกันน้อยกว่าภาพที่มีความแตกต่างกัน ดังเช่นรูปที่ 2.13 ดังนั้น ในโครงงานนี้ จึงได้ใช้ค่าความแตกต่างนี้มาเป็นตัววัดลักษณะของเฟรมภาพเพื่อทำการเปรียบเทียบกัน และนำไปสู่การพิจารณาตัดวิดีโออัตโนมัติในขั้นต่อไป

บทที่ 3

การออกแบบระบบ

3.1 เป้าหมายของการออกแบบโปรแกรม

เป้าหมายของโครงการคือ ออกแบบอัลกอริทึมและโปรแกรมที่จะใช้ตัดแบ่งภาพวิดีโอข่าวที่มีความยาวมากๆ ให้มีความยาวที่เล็กลง มีเนื้อหาข่าวที่เฉพาะเจาะจงเรื่องราว ทั้งนี้เพื่อสะดวกต่อการเก็บข้อมูล และง่ายต่อการเรียกใช้งาน เนื่องจากไฟล์ข่าวที่ได้หลังจากการประมวลผลของโปรแกรมที่ออกแบบนี้จะมีขนาดที่สั้นกะทัดรัด 1 ไฟล์จะมีเพียง 1 ข่าวเท่านั้น โดยผู้พัฒนาโครงงานได้อาศัย โครงสร้างโดยทั่วไปของภาพวิดีโอข่าว ซึ่งแสดงในรูปที่ 3.1

รูปที่ 3.1 โครงสร้างของวิดีโอข่าว

รูปที่ 3.1 วิดีโอข่าวโดยทั่วไปจะประกอบด้วยส่วนที่เป็นผู้ประกาศ คือผู้ที่แนะนำและอธิบายในส่วนของรายละเอียดต่างๆของข่าว และส่วนเนื้อหาของข่าวซึ่งก็คือภาพข่าวที่ถูกนำเสนอ เมื่อจบข่าวแล้วก็จะถูกตัดกลับมาที่ผู้ประกาศข่าวอีกครั้ง เพื่อนำเสนอข่าวต่อไปหรือไปสู่ช่วงของโฆษณา ผู้พัฒนาโครงการใช้หลักการนี้ในการออกแบบอัลกอริทึมที่จะให้ตัดแบ่งภาพวิดีโอข่าว นั่นคืออาศัยความแตกต่างของคุณสมบัติของภาพ ซึ่งก็คือ ภาพที่มีลักษณะและโทนสีใกล้เคียงกันจะมีคุณสมบัติใกล้เคียงกัน และในทางกลับกัน ภาพที่มีลักษณะแตกต่างกันมากๆ คุณสมบัติของสีก็ต่างกันด้วย ซึ่งขั้นตอนของการทำงานโปรแกรมจะถูกอธิบายในหัวข้อถัดไป

3.2 ลักษณะของข้อมูลเข้า (Input) และข้อมูลออก (Output)

ข้อมูลที่จะถูกป้อนให้กับโปรแกรมและผลลัพธ์สุดท้ายของโปรแกรมมีลักษณะดังนี้

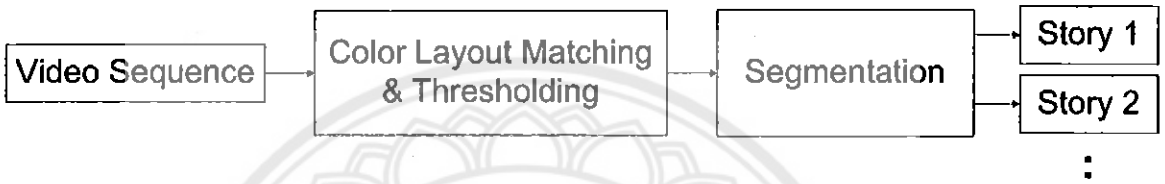
อินพุท : ภาพวิดีโอข่าวที่ถูกป้อนให้กับโปรแกรมจะมีลักษณะ เป็นวิดีโอข่าวที่มีความต่อเนื่องกัน และไฟล์ 1 ไฟล์จะประกอบไปด้วยข่าว 2 ข่าวขึ้นไปรวมกัน

เอาท์พุท : หลังจากการประมวลผลผู้ใช้จะได้ไฟล์ภาพวิดีโอข่าวหลายๆไฟล์ ซึ่งมีขนาดเล็กกว่าไฟล์ข่าวที่ถูกป้อนเข้ามาให้กับโปรแกรม แต่ละไฟล์ข่าวที่ได้จะมีเนื้อหาข่าวเฉพาะข่าวเดียวหรือเนื้อเรื่องเดียวเท่านั้น

ลักษณะอินพุทและเอาท์พุทที่ได้ของโปรแกรมแสดงในรูปที่ 3.2

Input Output

รูปที่ 3.2 แสดงลักษณะของอินพุทและเอาท์พุทของระบบ

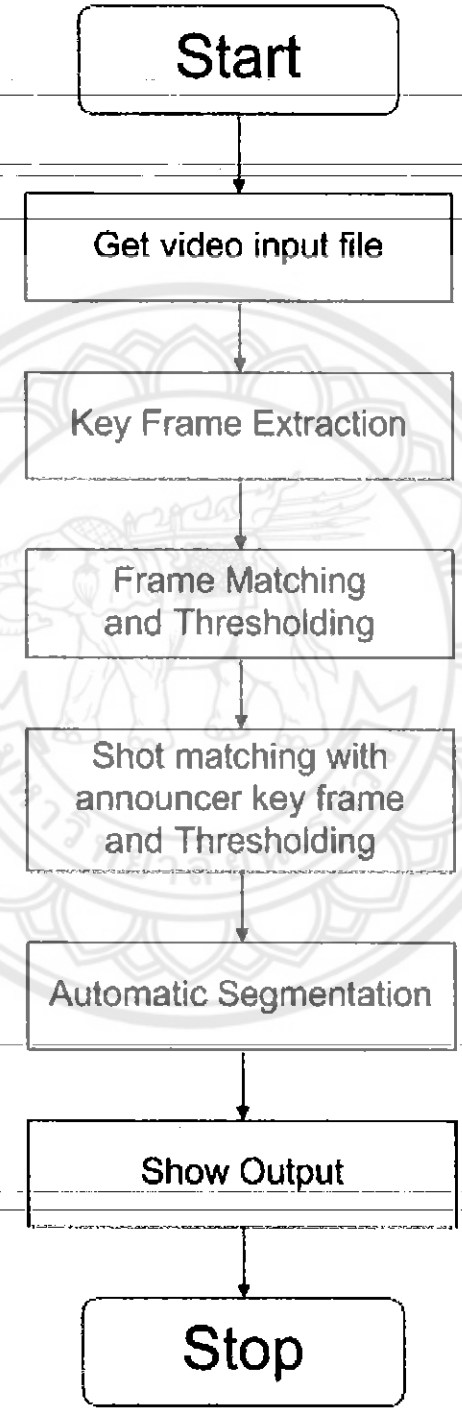


รูปที่ 3.3 แสดงลักษณะการทำงานโดยรวมของระบบ

รูปที่ 3.2 แสดงลักษณะของอินพุทที่ถูกป้อนเข้าไป ซึ่งเป็นวิดีโอที่มีหลายๆ**ข่าวต่อเนื่องกันไปจนจบ** และลักษณะของเอาท์พุทที่ได้หลังการประมวลผล ซึ่งการประมวลผล**โดยรวมของระบบจะเริ่มจาก**การ Matching สีของเฟรมที่ถูกเลือก เพื่อนำมาใช้พิจารณาในการตัด (Segmentation) **วิดีโอออก**มาเป็นเรื่องๆ

3.3 ผังงาน (Flowchart)

ขั้นตอนการดำเนินการของโปรแกรม ซึ่งได้ทำการอธิบายกระบวนการอย่างละเอียดในหัวข้อ 3.4 เป็นดังนี้



รูปที่ 3.4 ขั้นตอนการดำเนินการของโปรแกรม

3.4 วิธีดำเนินการ

จากขั้นตอนการดำเนินการของโปรแกรมในหัวข้อที่ 3.3 เริ่มต้นที่การรับข้อมูลอินพุท ซึ่งเป็นไฟล์วิดีโอข่าวเข้ามา โดยจะต้องเป็นไฟล์ที่ประกอบด้วยข่าวมากกว่า 2 ข่าวขึ้นไป จากนั้นนำข้อมูลอินพุทมาทำการประมวลผล ซึ่งในขั้นตอนนี้อาจแบ่งย่อยออกเป็น 2 ส่วน ได้แก่

1. การตัดแบ่งวิดีโอในระดับช็อต (Shot Segmentation)
2. การตัดแบ่งวิดีโอในระดับสตอรี่ (Story Segmentation)

ในส่วนแรกนั้น จะทำการตัดแบ่งไฟล์วิดีโอที่เป็นข้อมูลอินพุทออกเป็นช็อต ตามการวิเคราะห์ แต่ไม่ได้ตัดแบ่งออกเป็นไฟล์วิดีโอย่อย ๆ เลยทันที ต้องนำไปผ่านกระบวนการในส่วนที่ 2 เสียก่อน สำหรับในส่วนที่ 1 นี้ จะทำการนำเฟรมที่แตกออกมาจากวิดีโอข่าวมาใช้ในการประมวลผล ดังจะกล่าวเป็นตัวอย่างต่อไปในขั้นตอนการดำเนินการ

ในส่วนที่ 2 จะนำผลลัพธ์ของส่วนที่ 1 มาทำการวิเคราะห์ต่อ และมีการใช้ค่าเทรสโฮลด์ (Threshold) ในการวิเคราะห์เพื่อตัดแบ่งไฟล์วิดีโอข่าวออกเป็นเรื่อง ๆ เนื่องจากการตัดแบ่งวิดีโอในระดับเดียว อาจได้ผลลัพธ์ที่คลาดเคลื่อนไปมาก จึงต้องมีการตัดแบ่งเป็น 2 ระดับ โดยจะเห็นได้ชัดเจนจากการดำเนินการเป็นขั้นตอนโดยละเอียด ดังนี้

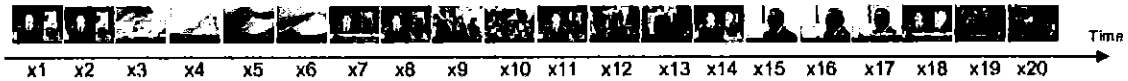
1. รับข้อมูลเข้าเป็นภาพวิดีโอข่าว โดยในรูปตัวอย่างด้านล่างเป็นไฟล์วิดีโอที่มีความยาวประมาณ 17 นาที อัตราเฟรม 30 FPS อัตราบิต (Bit rate) 365 kps ขนาดของภาพ (Dimension) 320 x 240

รูปที่ 3.5 ไฟล์วิดีโอตัวอย่างในการดำเนินการตามโปรแกรม

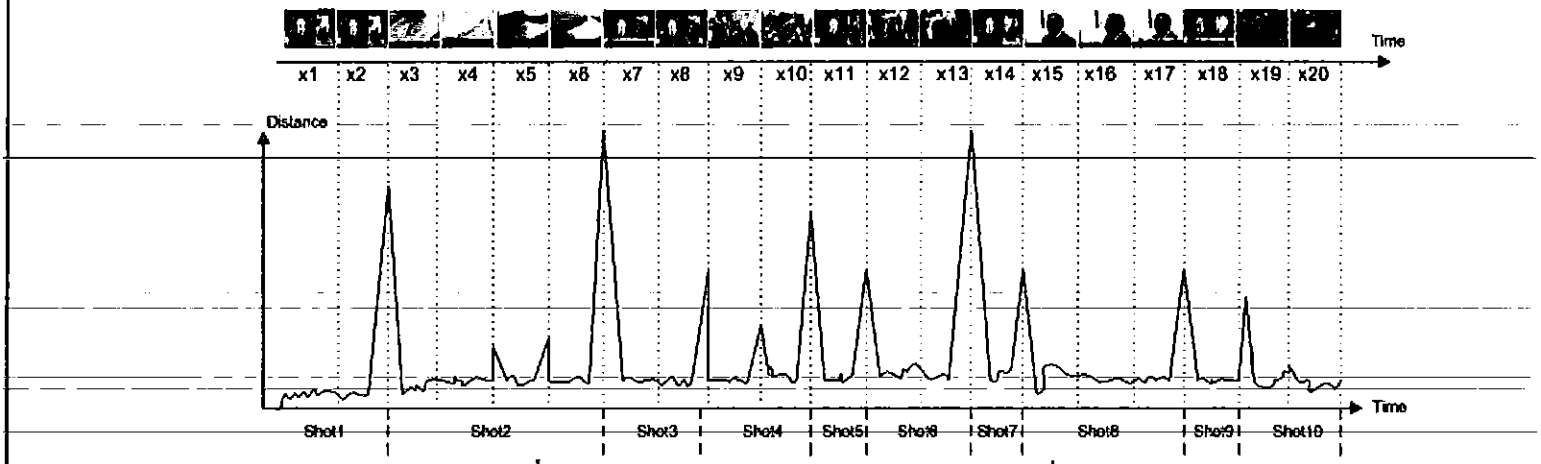
2. ทำการตัดแบ่งวิดีโอในระดับช็อต โดยเลือกเฟรมที่จะใช้ในการประมวลเพื่อเป็นคีย์เฟรม (Key Frame Extraction) ได้ดังนี้

รูปที่ 3.6 เฟรมที่เลือกมาเป็นคีย์เฟรมในระดับช็อต

3. ทำการแมทช์ชิ่งระหว่างเฟรมที่อยู่ติดกัน โดยใช้ทฤษฎีของการอธิบายภาพด้วยมาตรฐาน MPEG-7 โดยใช้กระบวนการของ Color Layout Descriptor ที่จะอธิบายภาพในระบบสี Y/Cb/Cr ทำการหาค่า Color Layout ของเฟรมภาพต่าง ๆ และนำมาเปรียบเทียบกันเพื่อหาค่าความแตกต่าง (Distance) ซึ่งจะใช้สมการที่ (2.1) จากบทที่ 2 ในการคำนวณความแตกต่างนี้ และจะได้กราฟความแตกต่างของเฟรม ดังรูปที่ 3.8



รูปที่ 3.7 เฟรมที่เลือกมาเรียงกันตามเวลา



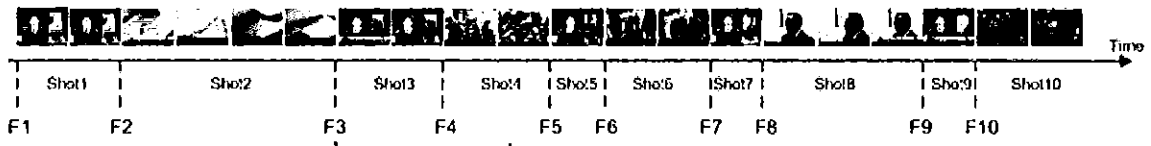
รูปที่ 3.8 กราฟแสดงความแตกต่างระหว่างเฟรมที่ติดกัน

รูปที่ 3.8 แสดงความแตกต่างระหว่างเฟรมที่ติดกันจากการหาค่าความแตกต่าง (Distance) ของ D(x1,x2) , D(x2,x3) , D(x3,x4) , ... , D(x19,x20)

ในการตัดแบ่งระดับช็อตนี้ ได้ทำการแบ่งจากเฟรมทั้ง 20 เฟรมได้ 10 ช็อต กราฟจะขึ้นสูงในระหว่างเฟรมที่มีความแตกต่างระหว่างกัน ยิ่งกราฟสูงมากก็แสดงให้เห็นว่ามีค่าความแตกต่างมาก ตัวอย่างเช่น ตำแหน่ง x6 กับ x7 มีค่าความแตกต่างระหว่างกันสูงมาก กราฟ ณ ตำแหน่งนี้ก็จะมีลักษณะสูงขึ้นไป แต่ตำแหน่ง x1 กับ x2 มีความคล้ายคลึงกัน กราฟที่ได้ก็จะมีความสูงไม่แตกต่างกันมากนัก โดยมีการเลือกใช้ค่าเทรสโฮลด์ (Threshold) ในการตัดแบ่งในระดับช็อตนี้

จะเห็นว่า การตัดแบ่งวิดีโอในระดับช็อตเพียงอย่างเดียวนี้ อาจไม่เพียงพอต่อการตัดแบ่งได้อย่างมีประสิทธิภาพ ตัวอย่างเช่นในตำแหน่ง x9 กับ x10 เป็นตำแหน่งที่อยู่ในข่าวเดียวกันคือเป็นเนื้อข่าวที่ต่อเนื่องกันอยู่ แต่ก็มีความแตกต่างระหว่างเฟรมทำให้กราฟขึ้นสูงไปเล็กน้อย หากทำการตัดแบ่งในส่วนนี้ไปเลยจะทำให้เนื้อหาของข่าวถูกตัดออกไป เราจึงต้องทำการวิเคราะห์เพื่อตัดแบ่งในอีกระดับหนึ่งเพื่อความแม่นยำ

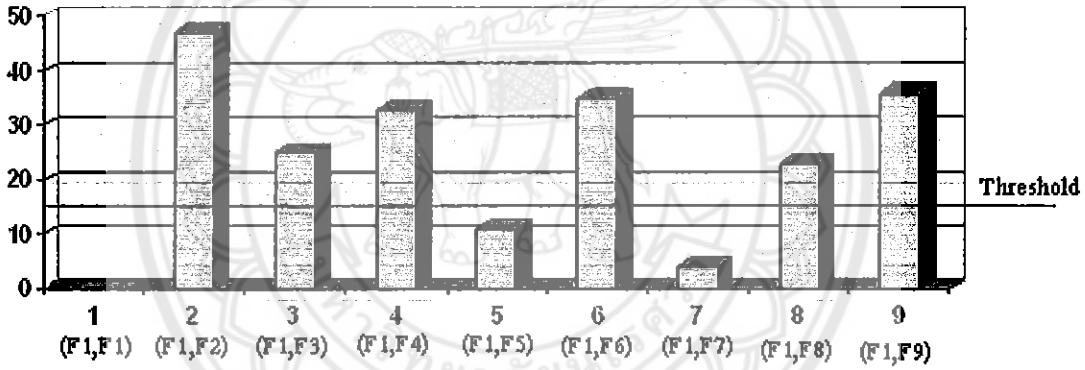
4. จากการตัดแบ่งระดับช็อตในข้อ 3. จะได้ช็อตจำนวน 10 ช็อต ทำการกำหนดคีย์เฟรมให้กับแต่ละช็อตที่ได้แบ่งออกมา



รูปที่ 3.7 คีย์เฟรมที่ได้จากการตัดแบ่งระดับช็อต

5. ในขั้นตอนนี้คือการตัดวิดีโอในระดับสตอรี่ โดยนำดีย์เฟรมที่กำหนดในข้อ 4. มาหาค่าความแตกต่างระหว่างกันอีกครั้งหนึ่ง โดยจะให้ดีย์เฟรมที่ 1 ในช็อตที่ 1 เป็นตัวหลักในการเปรียบเทียบระหว่างดีย์เฟรมต่าง ๆ เนื่องจากเป็นเฟรมของผู้ประกาศข่าว ทำให้ทราบว่าหากเฟรมใดที่มีค่าความแตกต่างกับ F1 น้อย หมายความว่ามีความแตกต่างระหว่างเฟรมน้อย นั่นคือน่าจะเป็นเฟรมของผู้ประกาศข่าวเช่นเดียวกับ F1 นั่นเอง โดยสมการที่ใช้ทำการหาค่าความแตกต่างยังคงเป็นสมการใน Color Layout Descriptor นั่นคือสมการที่ (2.1) เช่นเดิม

ในขั้นตอนนี้จะมีการกำหนดค่าเทรสโฮลด์ (Threshold) ที่เหมาะสมเพื่อเป็นจุดที่ช่วยตัดสินใจ หากค่าความแตกต่างในตำแหน่งใดที่ค่าต่ำกว่าค่าเทรสโฮลด์ (Threshold) ที่กำหนดไว้เนื่องจากน่าจะเป็นเฟรมของผู้ประกาศข่าวเช่นเดียวกัน หมายความว่า ณ ตำแหน่งที่ถูกนำไปเปรียบเทียบกับดีย์เฟรมที่ 1 นั้น เป็นเฟรมของผู้ประกาศข่าวอีกครั้งหนึ่ง นั่นคือได้ขึ้นเป็นข่าวเรื่องใหม่แล้ว ก็จะทำการตัดแบ่งวิดีโอ ณ ตำแหน่งที่อยู่ต่ำกว่าค่าเทรสโฮลด์ (Threshold) นั้น



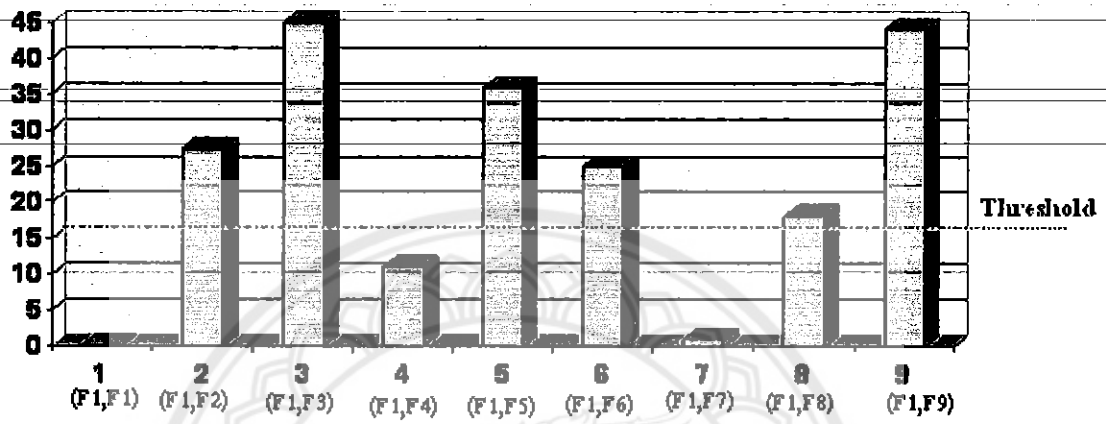
รูปที่ 3.8 กราฟแสดงความแตกต่างระหว่างดีย์เฟรมที่ได้

จากรูปที่ 3.8 สามารถอธิบายได้โดย

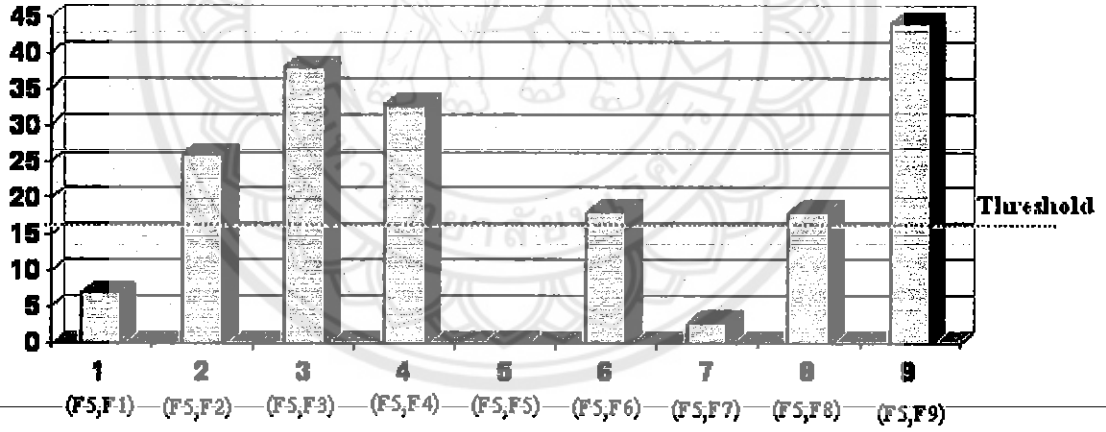
ตำแหน่งที่ 2 คือ การเปรียบเทียบความแตกต่างระหว่างดีย์เฟรมที่ 1 กับ 2 จะเห็นว่ามีความแตกต่างกันมาก เนื่องจากดีย์เฟรมที่ 1 เป็นส่วนของผู้ประกาศข่าว แต่ดีย์เฟรมที่ 2 เป็นส่วนของเนื้อหาข่าว ซึ่งมีการเปลี่ยนเฟรมจากผู้ประกาศข่าวไปเป็นเนื้อหาข่าว ทำให้เกิดความแตกต่างระหว่างเฟรมมาก ซึ่งตำแหน่งที่ 3 , 4 , 6 , 8 , 9 ก็มีผลลัพธ์ในทำนองเดียวกันนี้

ตำแหน่งที่ 1 คือ การเปรียบเทียบความแตกต่างระหว่างดีย์เฟรมที่ 1 กับ 1 จะเห็นว่าทั้งสองส่วนนี้ เป็นเฟรมของผู้ประกาศข่าวทั้งคู่ ทำให้มีความแตกต่างไม่มากนัก และมีค่าต่ำกว่าค่าเทรสโฮลด์ (Threshold) ดังนั้น โปรแกรมจะไม่ทำการตัดแบ่งในตำแหน่งที่มีค่าต่ำกว่าค่าเทรสโฮลด์ (Threshold) ซึ่งตำแหน่งที่ 5 , 7 ก็มีผลลัพธ์ในทำนองเดียวกันนี้

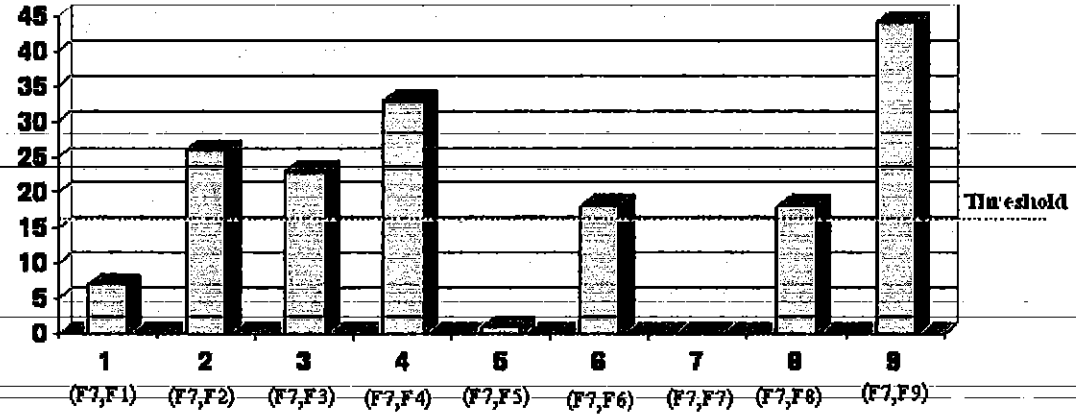
ในกรณีที่เฟรมของผู้ประกาศข่าวอาจมีหลายลักษณะ เพื่อให้ได้ความถูกต้องมากยิ่งขึ้น จึงได้มีการออกแบบให้มีการเพิ่มจำนวนเฟรมในการตัดระดับสตอรี่มากขึ้นเป็น 3 เฟรม โดยที่จะนำเฟรมทุกเฟรมหลักไปเปรียบเทียบกับทุกเฟรมที่ตัดได้ในระดับช็อต แล้วเลือกค่าความแตกต่างระหว่างเฟรมที่น้อยที่สุด



รูปที่ 3.9 กราฟแสดงความแตกต่างระหว่างคีย์เฟรมที่ 1 กับเฟรมที่ได้จากการตัดในระดับช็อต

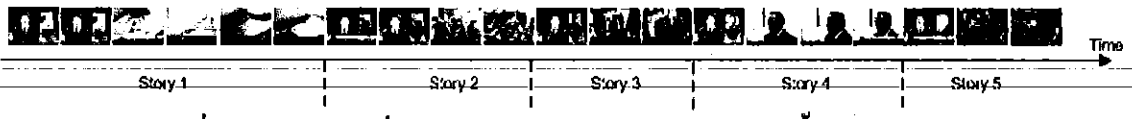


รูปที่ 3.10 กราฟแสดงความแตกต่างระหว่างคีย์เฟรมที่ 5 กับเฟรมที่ได้จากการตัดในระดับช็อต



รูปที่ 3.11 กราฟแสดงความแตกต่างระหว่างคีย์เฟรมที่ 7 กับเฟรมที่ได้จากการตัดในระดับช็อต

จากรูปที่ 3.9 จะเป็นการนำเฟรมที่หนึ่งที่เลือกมา มาเปรียบเทียบกับทุกๆเฟรมที่ได้จากการตัดในระดับช็อต ส่วนรูปที่ 3.10 และรูปที่ 3.11 แสดงการเปรียบเทียบระหว่างเฟรมที่ห้าและเจ็ดกับทุกๆ เฟรมที่ได้จากการตัดในระดับช็อต เมื่อทำการเปรียบเทียบครบทั้ง 3 เฟรมแล้วจะเลือกค่าความแตกต่างที่น้อยที่สุดไป เช่นในจุดที่ 1 การเปรียบเทียบโดยใช้เฟรมที่ (F1,F1) มีค่าความแตกต่างระหว่างเฟรมน้อยที่สุด เพราะฉะนั้นค่าความแตกต่างที่เลือกคือค่าความแตกต่างที่ได้จากการเปรียบเทียบโดยเฟรมที่ 1 จะมีค่าเท่ากับ 0 หรือในจุดที่ 3 จะพบว่าการเปรียบเทียบโดยใช้เฟรมที่ 7 (F7,F3) จะได้ค่าความแตกต่างน้อยที่สุดสำหรับเฟรมนี้จึงเลือกค่าความแตกต่างเป็น 23 คือค่าความแตกต่างที่ได้จากการเปรียบเทียบในเฟรมที่ 7 กับเฟรมที่ 3 นั่นเอง เมื่อได้ค่าความแตกต่างของทุกเฟรมแล้ว จะนำไปเปรียบเทียบว่าค่าความแตกต่างที่ได้มีค่าน้อยกว่าค่าเทรสโฮลด์ที่กำหนดหรือไม่ ถ้ามีค่าน้อยกว่าแสดงว่าเฟรมนั้นมีการเปลี่ยนการนำเสนอข่าวหรือจบข่าวแล้ว จะมีการเก็บเฟรมนั้นไว้เพื่อใช้ระบุตำแหน่งในการทำการตัดวิดีโอต่อไป จากรูปที่ 3.9 , 3.10 และ 3.11 จะเห็นว่าจะมีการตัดวิดีโอที่ตำแหน่งเฟรมที่ 1 ถึงเฟรมที่ 5 , เฟรมที่ 5 ถึงเฟรมที่ 7 และเฟรมที่ 7 ถึงเฟรมที่ 9 เป็นต้น

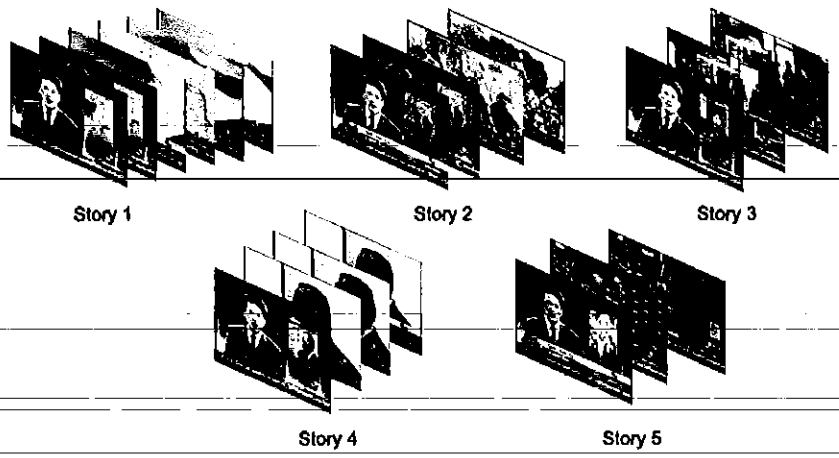


รูปที่ 3.12 ผลลัพธ์ที่ได้จากการตัดแบ่งไฟล์ภาพวิดีโอในขั้นสุดท้าย

จะเห็นว่ามีการตัดแบ่งข่าวออกเป็น 5 เรื่อง ดังรูปที่ 3.13

ห้องสมุดคณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร

149788
2/5.
44847
2551



รูปที่ 3.13 ไฟล์ภาพวิดีโอ 5 เรื่องจากการตัดแบ่ง

ในการทำงานของโปรแกรมอาจจะมีความผิดพลาดหรือคลาดเคลื่อนไปบ้าง ซึ่งสามารถตรวจสอบความถูกต้องได้ในหัวข้อของการวิเคราะห์ความถูกต้องของผลลัพธ์

3.5 การวิเคราะห์ความถูกต้อง

การวิเคราะห์ความถูกต้องในการตัดวิดีโอ พิจารณาผลลัพธ์สุดท้ายที่ได้จากโปรแกรม โดยจะพิจารณาจากวิดีโอเอาท์พุทที่ได้ ว่ามีความถูกต้องในการตัดเนื้อหาของข่าวได้ตามจุดประสงค์หรือไม่ ซึ่งก็คือการตัดข่าวออกมาเป็นเรื่องๆ โดยมีเนื้อหาข่าวที่เฉพาะเจาะจง และมีเนื้อหาที่ครบถ้วนตามอินพุทที่ได้รับมา ไม่ตัดในวิดีโอในขณะที่ยังนำเสนอข่าวไม่จบเป็นต้น ซึ่งการวัดประสิทธิภาพในการตัดวิดีโอในระดับนี้ วัดได้ดังสมการ

$$precision(\%) = \frac{N_R}{N_T} \times 100 \quad (3.1)$$

สมการข้างต้นเป็นการวัดความถูกต้องในการตัดแบ่งวิดีโอของโปรแกรม ซึ่งจะให้ค่าออกมาเป็นเปอร์เซ็นต์ความถูกต้อง โดยที่ N_R คือ จำนวนของข่าวที่โปรแกรมตัดได้ ซึ่งจะนำไปเทียบค่า N_T ซึ่งก็คือจำนวนข่าวที่ได้โดยไม่ใช้โปรแกรม

3.6 เครื่องมือที่ใช้พัฒนาโครงงาน

โปรแกรมใช้ทางผู้จัดทำใช้พัฒนาคือ Microsoft Visual C# .Net โดยใช้แม่แบบเป็น Window application

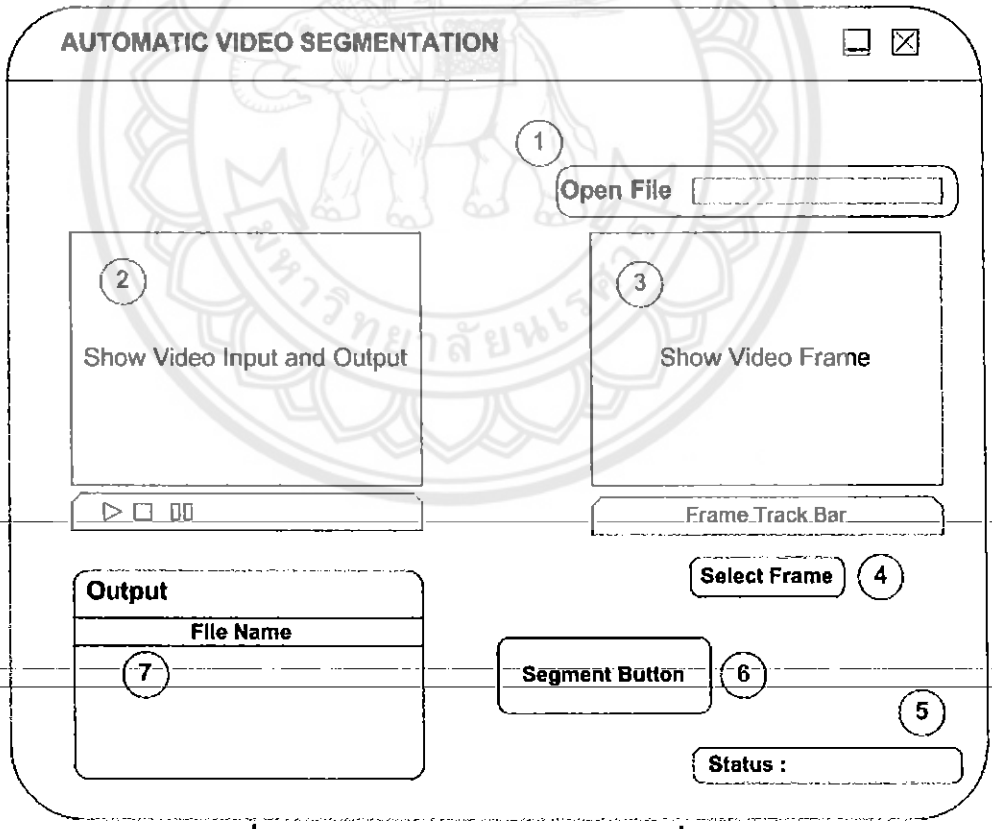
ซีชาร์ป (C#) เป็นภาษาโปรแกรมเชิงวัตถุ (Object-oriented programming) ทำงานบนดอตเน็ตเฟรมเวิร์ก ซึ่งมีรากฐานมาจากภาษาซีพลัสพลัสและจาวา ในการทำงานจะมีการมองทุก

อย่างเป็นวัตถุ และสืบทอดต่อกันมา ด้วยการสร้างคลาส คลาสทุกคลาสที่ไม่ถูกสืบทอดจากคลาสอื่นจะถือว่าคลาสนั้นสืบทอดมาจากคลาสชื่อ object ซึ่งก็คือคลาสทุกคลาสที่ถูกสร้างมา จะถือว่าเป็นวัตถุอย่างหนึ่ง และเมื่อนำคลาสนั้นไปให้คลาสอื่นสืบทอด คลาสที่ได้รับการสืบทอดก็จะสืบทอดความเป็น object มาด้วย

ในภาษาซีชาร์ปจะไม่มีการใช้ไฟล์ .h แต่จะใช้ไฟล์ .dll ซึ่งสร้างขึ้นบนดอตเน็ตเฟรมเวิร์ก ซึ่ง .dll ชนิดนี้ จะมีโครงสร้างส่วนบนสุดของไฟล์เป็นโครงสร้างของคลาสต่างๆที่เก็บอยู่ในไฟล์นั้นๆ นั่นคือ ไลบรารีบนดอตเน็ตเฟรมเวิร์ก มีการเก็บ .h รวมกับ .dll และสามารถเรียกใช้ในการเขียนโปรแกรมได้ทันที เมื่อมีการเพิ่มไฟล์ .dll เข้าไปในโปรเจค

3.7 การออกแบบหน้าตาของโปรแกรม

ลักษณะรูปแบบของโปรแกรมที่ออกแบบ เป็นดังนี้



รูปที่ 3.14 ลักษณะรูปแบบของโปรแกรมที่ออกแบบ

1. ส่วนของการ Open File เป็นการให้ผู้ใช้เปิดไฟล์วิดีโอที่ต้องการตัดแบ่งขึ้นมาจากโฟลเดอร์ใด ๆ ที่ผู้ใช้ได้เก็บไฟล์วิดีโอไว้

2. ส่วนของหน้าจอแสดงภาพวิดีโอ มีปุ่มเล่นไฟล์วิดีโอ , หยุดไฟล์วิดีโอ และหยุดการเล่นไฟล์วิดีโอชั่วคราว

3. ส่วนของหน้าจอแสดงเฟรมภาพวิดีโอที่ทำการแตกออกเป็นเฟรมภาพแล้ว

4. ปุ่ม Select Frame เป็นปุ่มกดเลือกเฟรมเพื่อนำไปเป็นคีย์เฟรมใช้เปรียบเทียบในระดับสตอรี่ ซึ่งจะต้องเลือกให้เป็นเฟรมของผู้ประกาศข่าว

5. แถบ Status เป็นแถบแสดงการทำงานของโปรแกรม

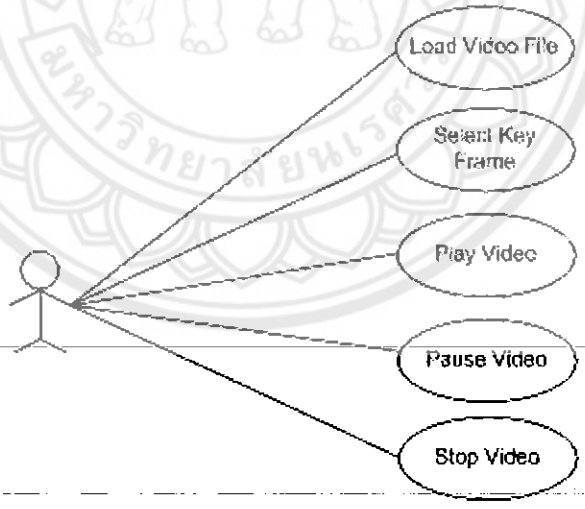
6. ส่วนของปุ่มกด Segment คือ กดเลือกว่าให้ทำการตัดแบ่งไฟล์วิดีโอต้นฉบับที่ได้ทำการเลือกมาแล้ว (Segment Button)

7. ส่วนของ Output แสดงเป็นตารางของไฟล์วิดีโอที่ได้ทำการตัดแบ่งแล้ว ซึ่งอาจมีหลายไฟล์ แล้วแต่ไฟล์วิดีโอต้นฉบับ สามารถทำการเล่นไฟล์วิดีโอเพื่อดูผลลัพธ์ที่ออกมาได้

3.8 การออกแบบโปรแกรม

ไดอะแกรมต่าง ๆ ในการออกแบบการพัฒนาโปรแกรมโดยใช้ Microsoft Visual C# .Net เป็นดังนี้

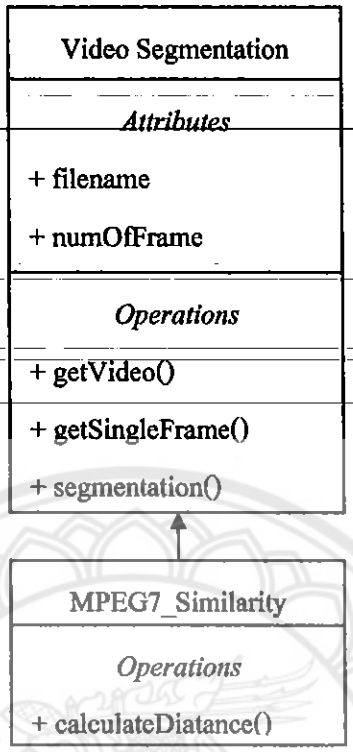
3.8.1 Use Case Diagram



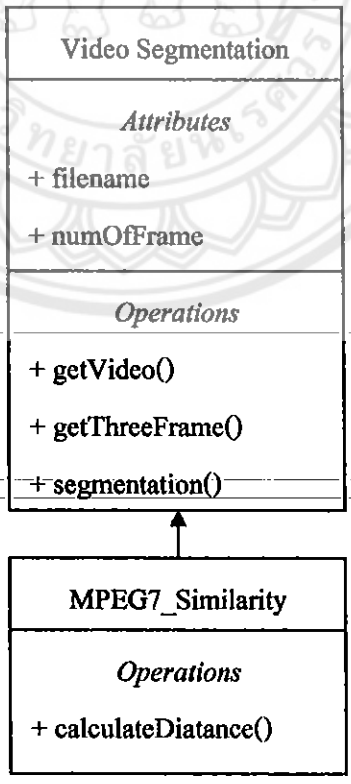
รูปที่ 3.15 Use Case Diagram

จากรูปที่ 3.15 แสดง Use Case Diagram ซึ่งเป็นส่วนแสดงแบบจำลองเพื่อใช้อธิบายหน้าที่ของโปรแกรมที่ผู้ใช้สามารถใช้งานได้

3.8.2 Class Diagram



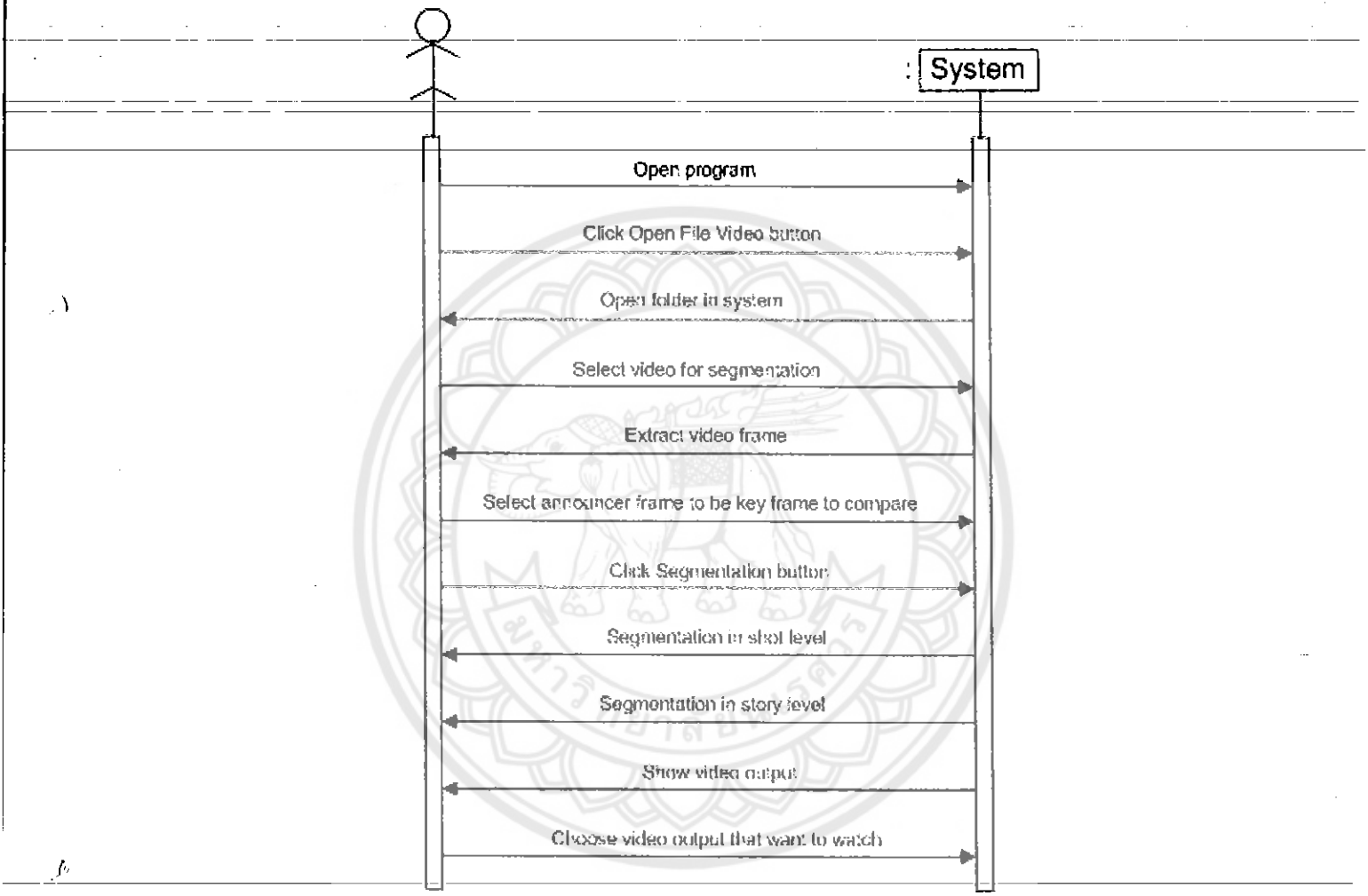
รูปที่ 3.16 Class Diagram สำหรับให้ผู้ใช้เลือก 1 เฟรมในการเปรียบเทียบระดับสตอรี่



รูปที่ 3.17 Class Diagram สำหรับให้ผู้ใช้เลือก 3 เฟรมในการเปรียบเทียบระดับสตอรี่

จากรูปที่ 3.16 และ 3.17 เป็น Class Diagram ประกอบด้วยคลาส , ความสัมพันธ์ต่างๆ ระหว่างคลาส และแสดงรายละเอียดภายในคลาส

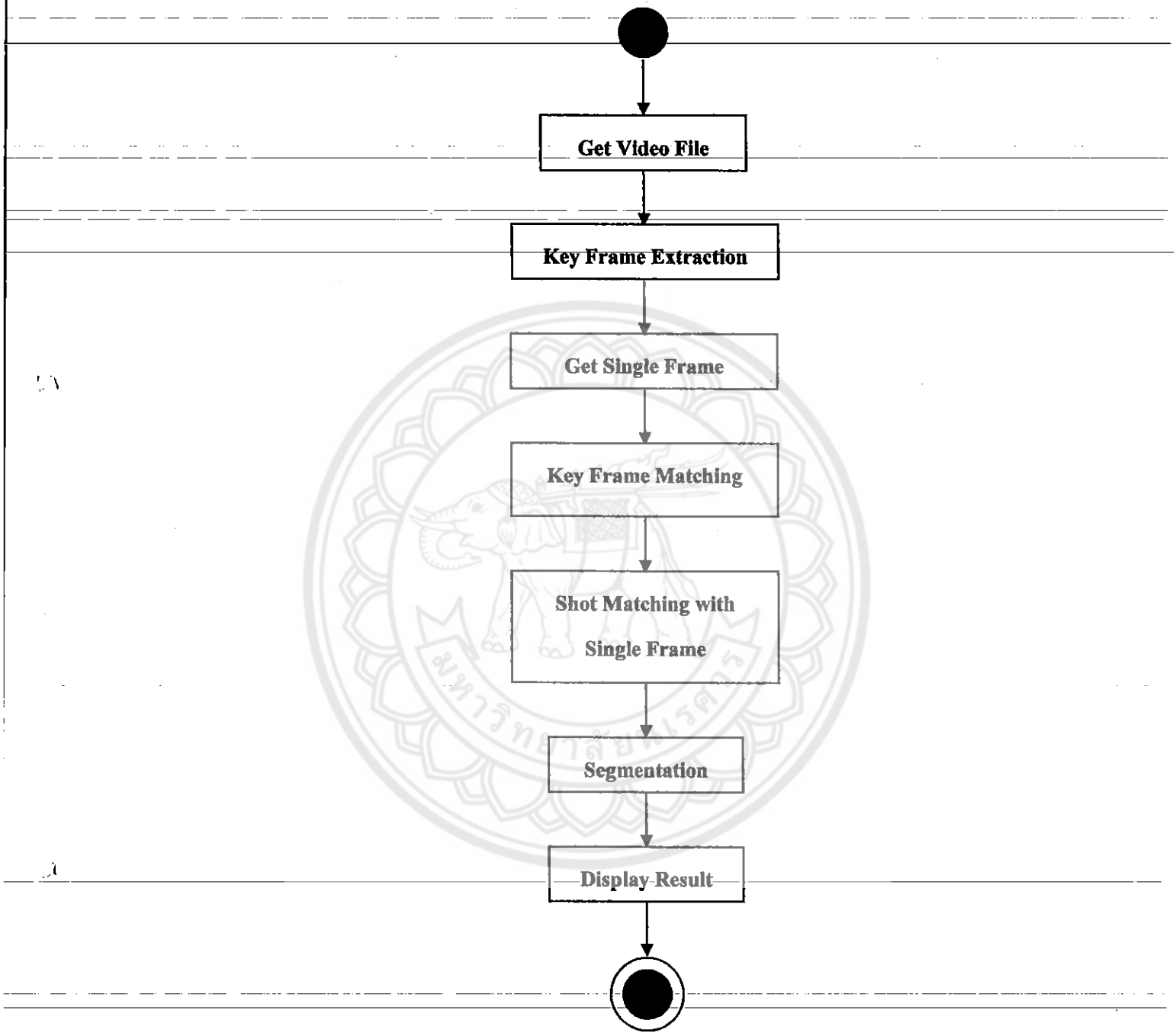
3.8.3 Sequence Diagram



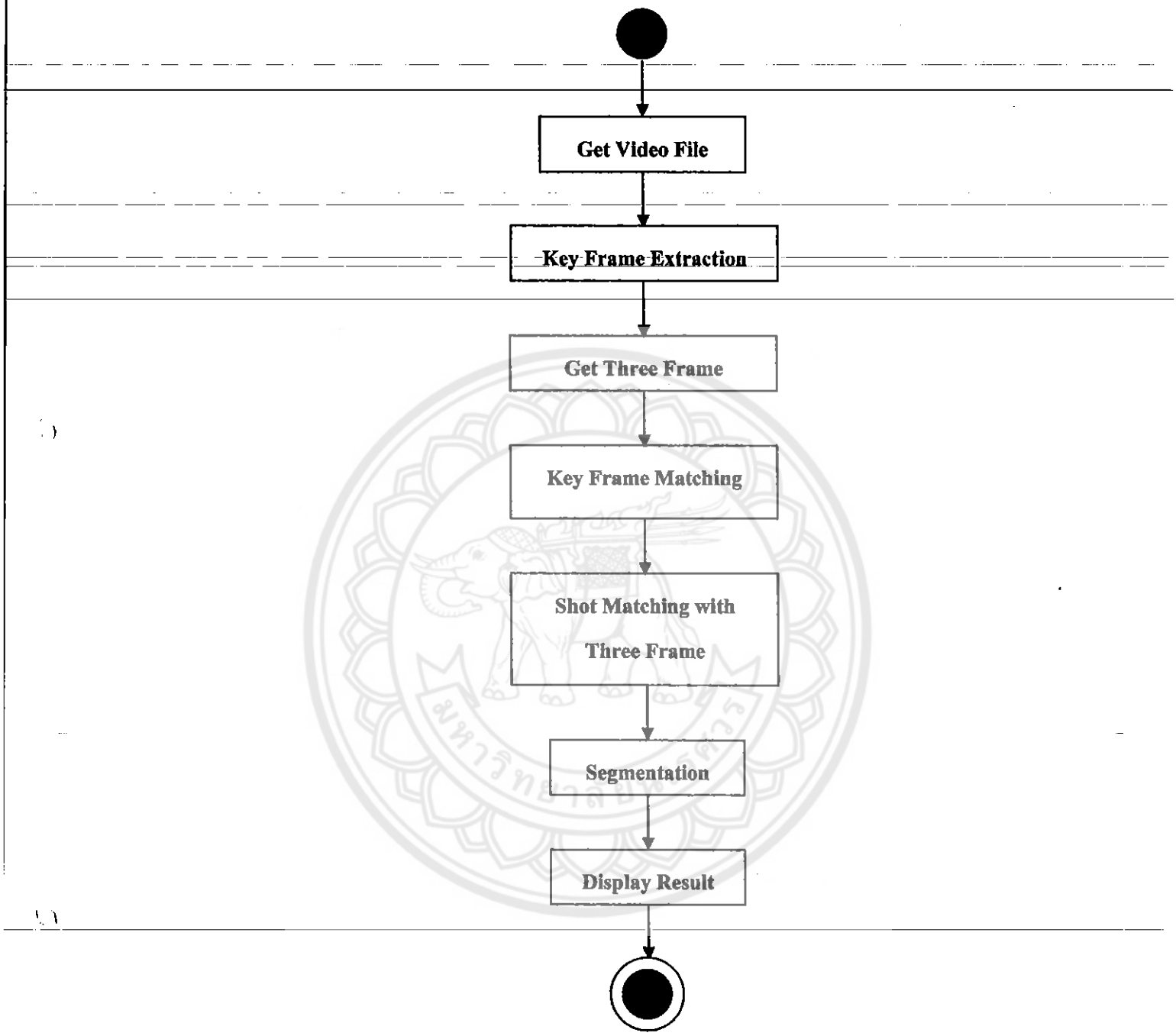
รูปที่ 3.18 Sequence Diagram

จากรูปที่ 3.18 แสดง Sequence Diagram ซึ่งเป็นไดอะแกรมที่แสดงลำดับการทำงานของระบบ ที่มีความสัมพันธ์ระหว่างระหว่างผู้ใช้และระบบ

3.8.4 Activity Diagram



รูปที่ 3.19 Activity Diagram สำหรับให้ผู้ใช้เลือก 1 เฟรมในการเปรียบเทียบระดับสตอรี่



รูปที่ 3.20 Activity Diagram สำหรับให้ผู้ใช้เลือก 3 เฟรมในการเปรียบเทียบระดับสตอรี่

รูปที่ 3.19 และ 3.20 คือ Activity Diagram แสดงลำดับกิจกรรมของการทำงานของระบบ และขั้นตอนการทำงานในการปฏิบัติการ

บทที่ 4

การดำเนินงานและผลการทดลอง

4.1 ข้อมูลเบื้องต้นของไฟล์วิดีโอข่าว

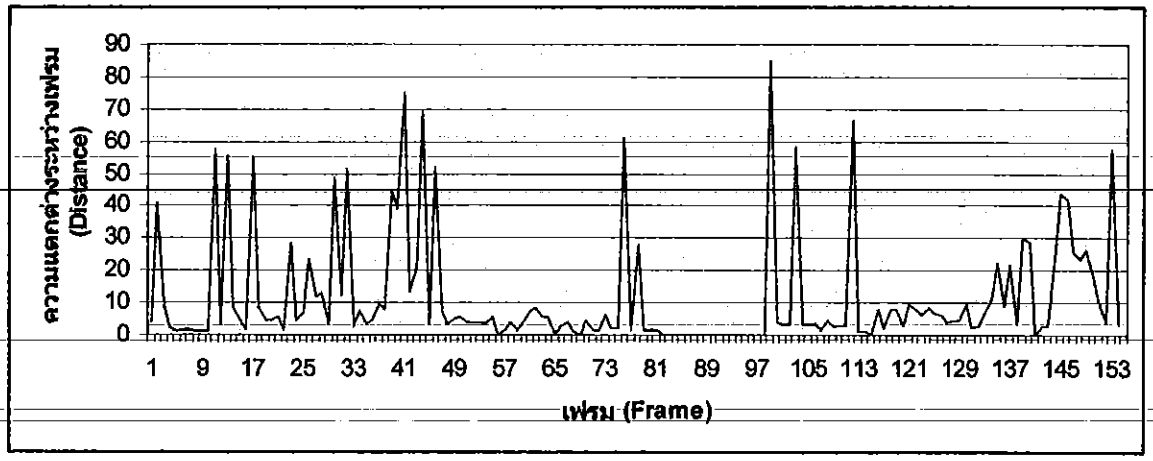
ไฟล์วิดีโอข่าวที่นำมาทดสอบโปรแกรมมีจำนวน 9 ไฟล์ โดยทำการเก็บข้อมูลและรายละเอียดของไฟล์วิดีโอข่าวด้วยการเปิดดูในโปรแกรมแสดงวิดีโอด้วยตัวผู้จัดทำ เพื่อนำไปเปรียบเทียบกับผลลัพธ์ที่ได้จากผลการรันโปรแกรม

ตารางที่ 4.1 รายละเอียดของไฟล์วิดีโอข่าว

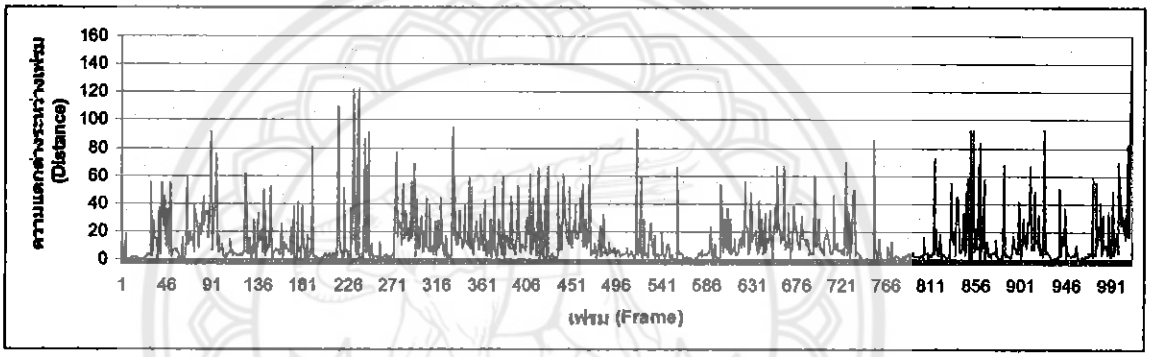
ชื่อไฟล์	จำนวนเฟรมที่แตกออกมาได้	วัดด้วยสายตา	
		จำนวนช็อต	จำนวนข่าว
news_nbt_01	273	51	3
news_nbt_02	161	32	2
news_ch3_01	155	22	3
news_ch3_02	368	74	5
news_ch3_03	522	71	3
news_modern-9_01	1011	122	6
news_modern-9_02	534	86	4
news_modern-9_03	358	73	4
news_modern-9_04	428	73	7

4.2 ผลการทดลองตัดแบ่งวิดีโอข่าว

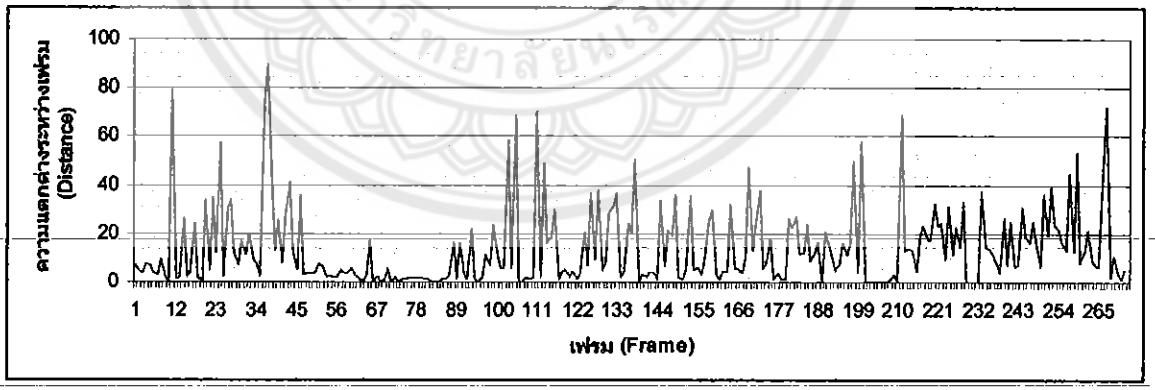
ทำการทดลองตัดแบ่งวิดีโอข่าว ซึ่งเป็นการตัดวิดีโอในขั้นแรก ค่าเทรสโฮลด์ ระดับช็อต คือ 20 โดยวิเคราะห์จากตัวอย่างกราฟของความสัมพันธ์ระหว่างเฟรมที่อยู่ติดกันกับค่าความแตกต่าง (Distance) ที่ใช้สมการการจับคู่ (Matching) ของ Color Layout ดังนี้



รูปที่ 4.1 กราฟความสัมพันธ์ระหว่างเฟรมกับค่าความแตกต่างของไฟล์วิดีโอ news_ch3_01



รูปที่ 4.2 กราฟความสัมพันธ์ระหว่างเฟรมกับค่าความแตกต่างของไฟล์วิดีโอ news_modern-9_01



รูปที่ 4.3 กราฟความสัมพันธ์ระหว่างเฟรมกับค่าความแตกต่างของไฟล์วิดีโอ news_nbt_01

จากรูปที่ 4.1 , 4.2 และ 4.3 ซึ่งเป็นกราฟแสดงความสัมพันธ์ระหว่างเฟรมกับค่าความแตกต่างของไฟล์วิดีโอตัวอย่าง ทำให้สามารถตัดสินใจเลือกค่า 20 มาเป็นค่าเทรสโฮลด์ในการตัดแบ่งระดับช็อตได้ เนื่องจากค่านี้มีความละเอียดพอที่จะเก็บรายละเอียดของช็อตย่อย ๆ ในไฟล์วิดีโอ เพื่อนำไปเปรียบเทียบในระดับสตอรี่อีกครั้งหนึ่งได้

สำหรับการตัดแบ่งวิดีโอข่าวในระดับสตอรี่นั้น จะทำการเลือกคีย์เฟรมของผู้ประกาศข่าวที่หน้าต่างของโปรแกรม เพื่อที่โปรแกรมจะนำคีย์เฟรมนั้นมาทำการเปรียบเทียบกับช็อตต่าง ๆ ที่ตัดไว้อีกครั้งหนึ่ง ถ้าพบว่าค่าความแตกต่างระหว่างคีย์เฟรมที่เลือกกับเฟรมที่ตัดในระดับช็อตมีน้อย แสดงว่าทั้งสองเฟรมนั้นอาจจะเป็นเฟรมของผู้ประกาศข่าวเหมือนกัน ซึ่งแต่ละไฟล์วิดีโอนั้นค่าความแตกต่างที่สมควรจะใช้ตัดแบ่งนั้นมีค่าต่าง ๆ กันไป ตามแต่รายละเอียดของวิดีโอข่าวนั้น ๆ จึงต้องทำการเปรียบเทียบค่าที่จะใช้ตัดในระดับสตอรี่ค่าต่าง ๆ กับวิดีโอข่าวไฟล์ต่าง ๆ เพื่อหาระดับของค่าเทรสโฮลด์ที่เหมาะสมที่สุด

ทั้งนี้ ข้อมูลก่อนรูปตัวอย่างของวิดีโอข่าว คือข้อมูลตามความเป็นจริง ส่วนข้อมูลในตาราง คือข้อมูลที่ได้จากโปรแกรม โดยรายละเอียดของผลลัพธ์ที่ได้ในแต่ละไฟล์วิดีโอที่ใช้ค่าเทรสโฮลด์ในระดับสตอรี่ค่าต่าง ๆ มีดังนี้

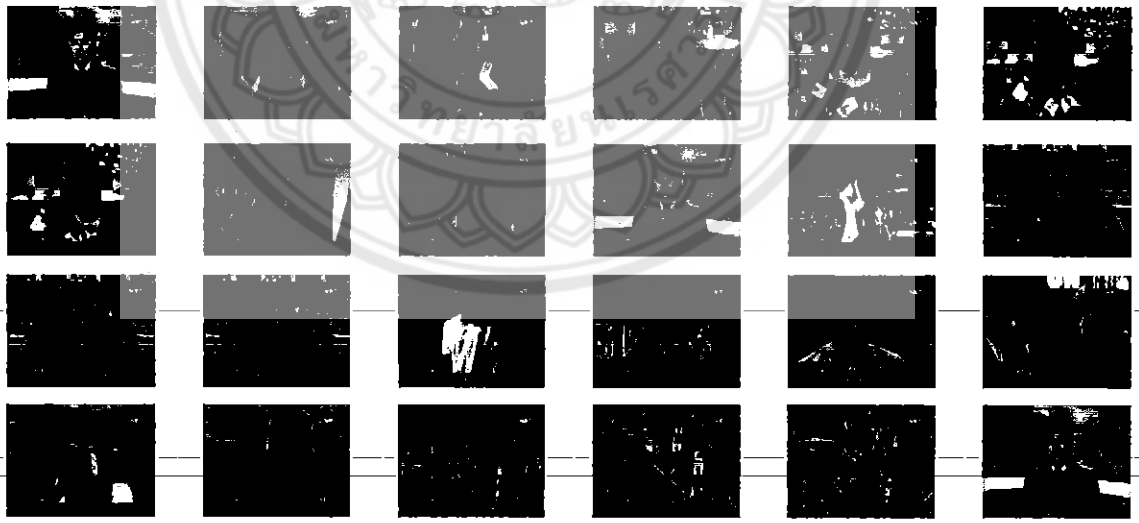
1. ไฟล์วิดีโอข่าว news_nbt_01

ชื่อไฟล์ : news_nbt_01 *ความยาววิดีโอ (วินาที) : 273*

จำนวนเฟรมต่อวินาที : 30 *จำนวนเฟรมทั้งหมด : 8190*

จำนวนช็อต : 51 *จำนวนข่าว : 3*

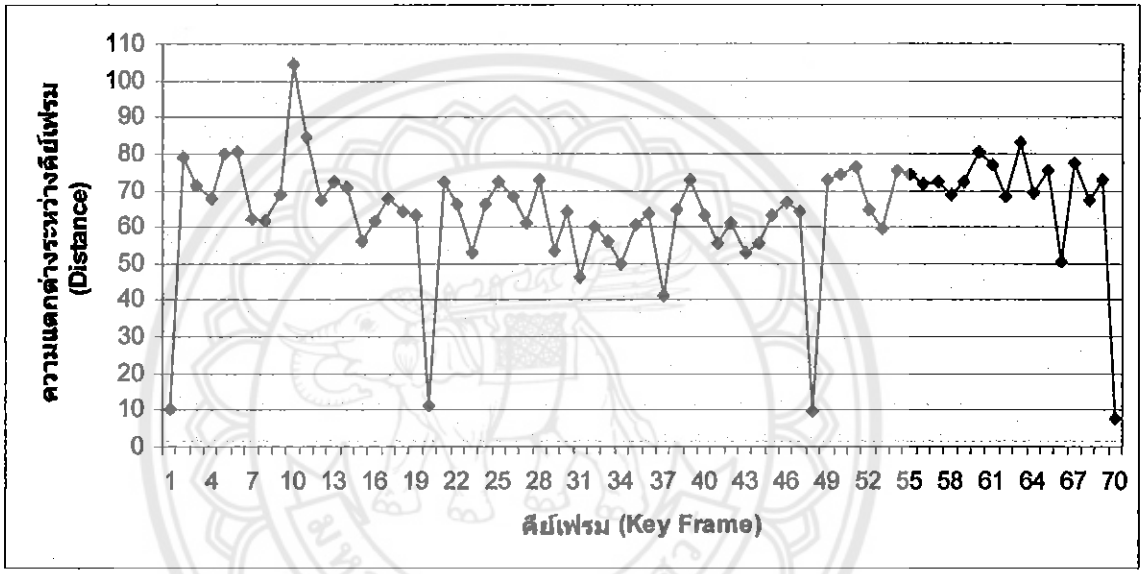
ช่วงเวลาของแต่ละข่าว (วินาทีที่) : 1-105 , 106-200 , 201-273



รูปที่ 4.4 ตัวอย่างภาพในไฟล์วิดีโอข่าว news_nbt_01

ตารางที่ 4.2 ผลลัพธ์การตัดแบ่งวิดีโอในเทรสโฮลด์ระดับต่าง ๆ ของไฟล์วิดีโอ news_nbt_01

ค่าเทรสโฮลด์ ระดับฮตอรี่	จำนวนข่าว ที่โปรแกรมตัด	ช่วงเวลาของแต่ละข่าว (วินาทีที่)
5	1	1-273
10	3	1- 105 , 106 - 200 , 201-273
13	3	1- 105 , 106 - 200 , 201-273
15	4	1-105 , 106-200 , 201-268 , 268-273



รูปที่ 4.5 กราฟความสัมพันธ์ระหว่างคีย์เฟรมของผู้ประกาศข่าวกับคีย์เฟรมที่ตัดในระดับช็อตของไฟล์วิดีโอ news_nbt_01

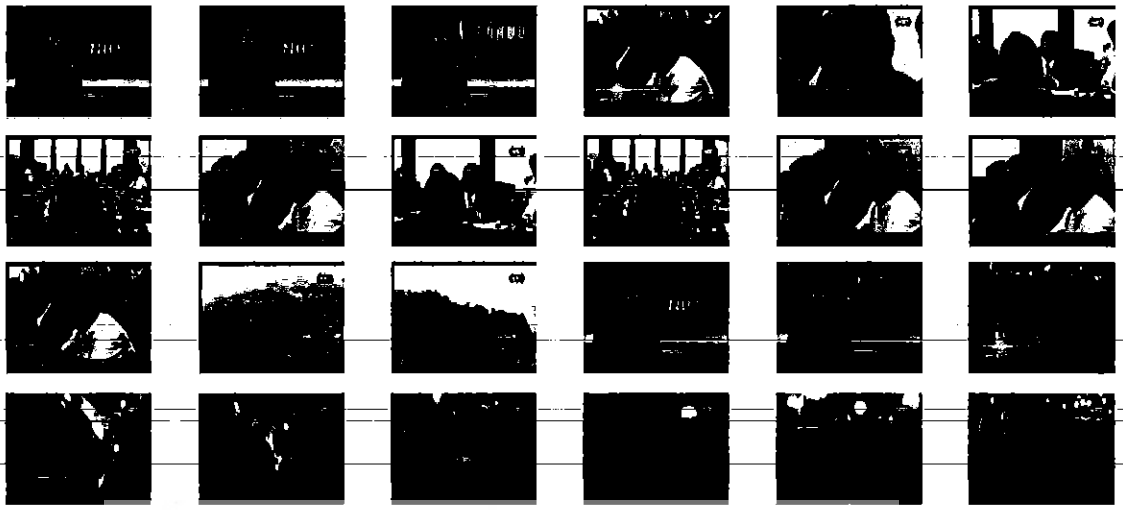
2. ไฟล์วิดีโอข่าว news_nbt_02

ชื่อไฟล์ : news_nbt_02 *ความยาววิดีโอ(วินาที) : 161*

จำนวนเฟรมต่อวินาที : 30 *จำนวนเฟรมทั้งหมด : 4830*

จำนวนช็อต : 32 *จำนวนข่าว : 2*

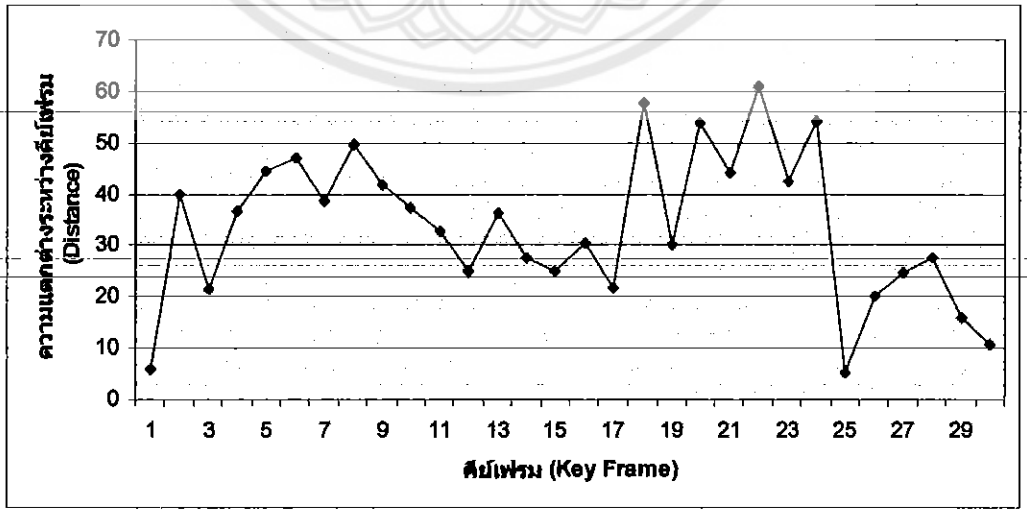
ช่วงเวลาของแต่ละข่าว (วินาทีที่) : 1-94 , 95-161



รูปที่ 4.6 ตัวอย่างภาพในไฟล์วิดีโอข่าว news_nbt_02

ตารางที่ 4.3 ผลลัพธ์การตัดแบ่งวิดีโอในเทรสโฮลด์ระดับต่าง ๆ ของไฟล์วิดีโอ news_nbt_02

ค่าเทรสโฮลด์ ระดับสตอรี่	จำนวนข่าว ที่โปรแกรมตัด	ช่วงเวลาของแต่ละข่าว (วินาทีที่)
5	2	1-94 , 95-161
10	2	1-94 , 95-161
13	3	1-94 , 95-156 , 157-161
15	3	1-94 , 95-156 , 157-161
20	5	1-94 , 95-106, 107-148 , 149-156 , 157-161



รูปที่ 4.7 กราฟความสัมพันธ์ระหว่างคีย์เฟรมของผู้ประกาศข่าวกับคีย์เฟรมที่ตัดในระดับช็อตของไฟล์วิดีโอ news_nbt_02

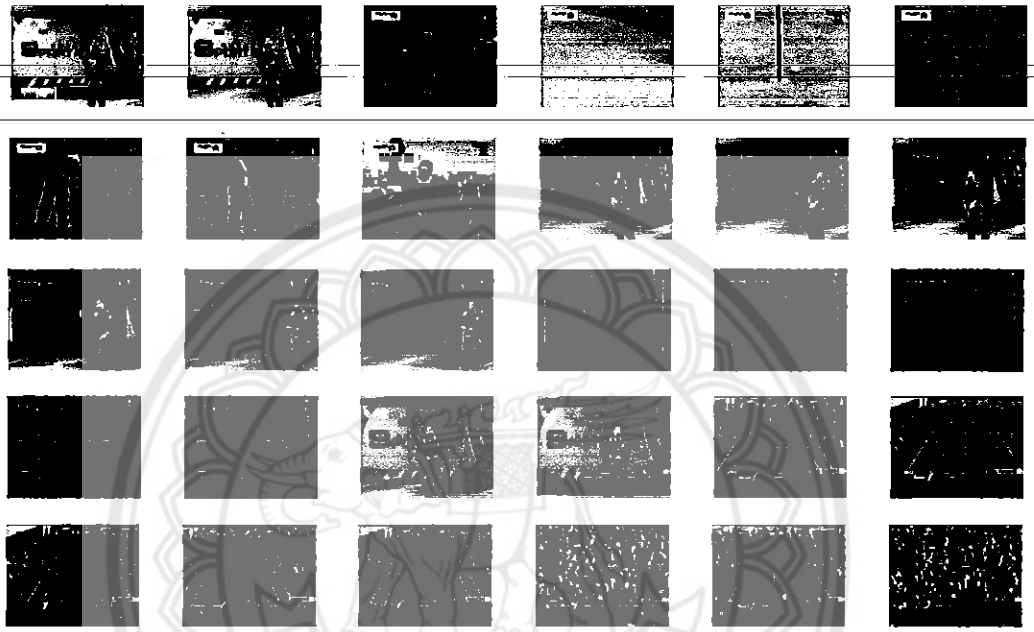
3. ไฟล์วิดีโอข่าว news_ch3_01

ชื่อไฟล์ : news_ch3_01 *ความยาววิดีโอ (วินาที) : 155*

จำนวนเฟรมต่อวินาที : 29 *จำนวนเฟรมทั้งหมด : 4495*

จำนวนช็อต : 21 *จำนวนข่าว : 3*

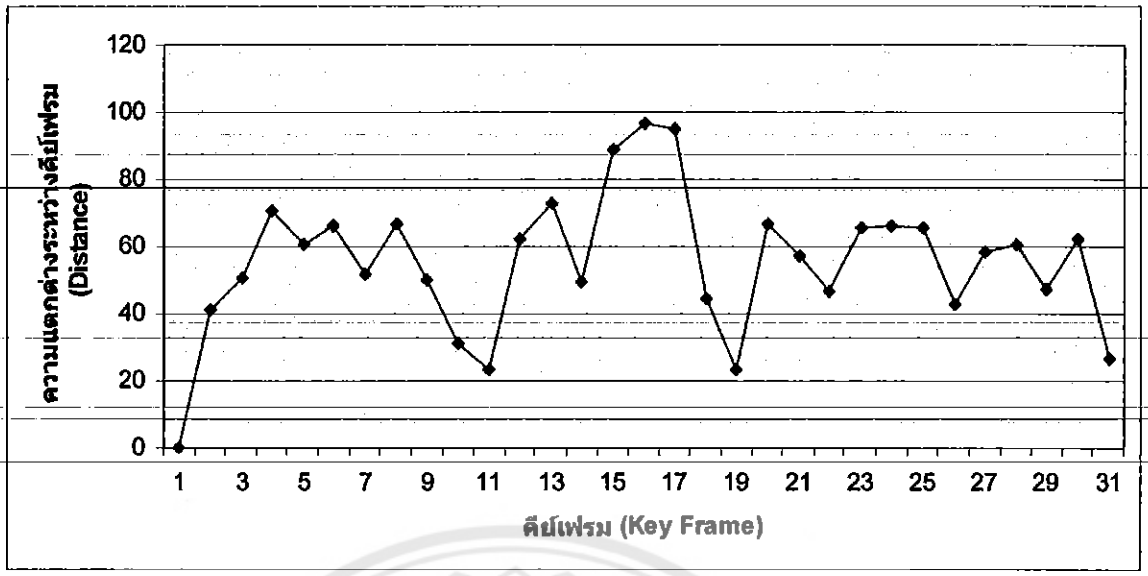
ช่วงเวลาของแต่ละข่าว (วินาทีที่) : 1-44 , 45-139 , 140-155



รูปที่ 4.8 ตัวอย่างภาพในไฟล์วิดีโอข่าว news_ch3_01

ตารางที่ 4.4 ผลลัพธ์การตัดแบ่งวิดีโอในเทรสโฮลด์ระดับต่าง ๆ ของไฟล์วิดีโอ news_ch3_01

ค่าเทรสโฮลด์ ระดับสตอรี่	จำนวนข่าว ที่โปรแกรมตัด	ช่วงเวลาของแต่ละข่าว (วินาทีที่)
15	1	1 - 155
17	1	1 - 155
20	1	1 - 155
25	3	1-40 , 41-103 , 104-155
27	4	1-40 , 41-103 , 104-153 , 154-155



รูปที่ 4.9 กราฟความสัมพันธ์ระหว่างคีย์เฟรมของผู้ประกาศข่าวกับคีย์เฟรมที่ตัดในระดับช็อตของไฟล์วิดีโอ news_ch3_01

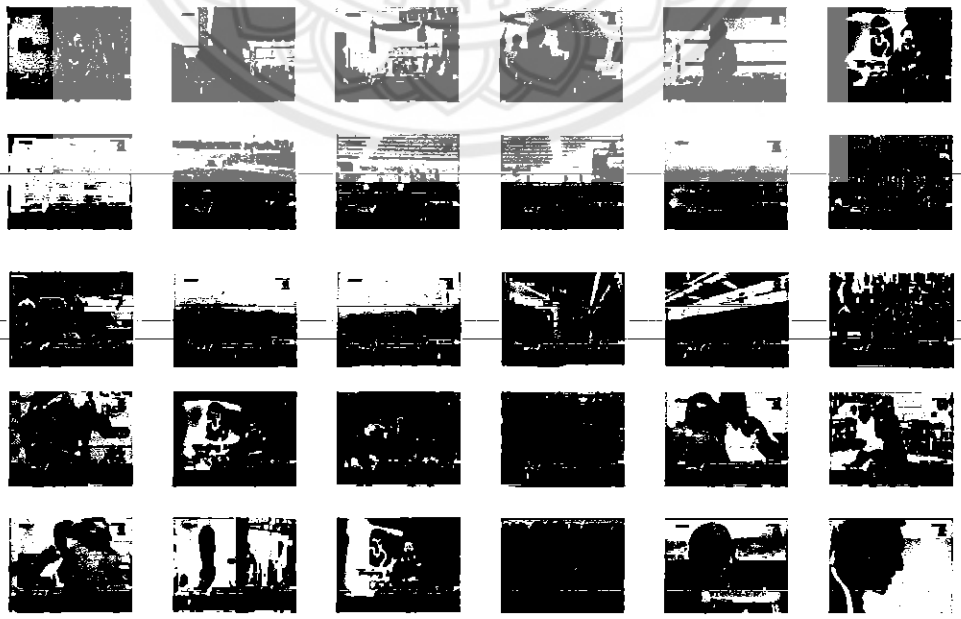
4. ไฟล์วิดีโอข่าว news_ch3_02

ชื่อไฟล์ : news_ch3_02 *ความยาววิดีโอ(วินาที) : 368*

จำนวนเฟรมต่อวินาที : 30 *จำนวนเฟรมทั้งหมด : 11040*

จำนวนช็อต : 74 *จำนวนข่าว : 5*

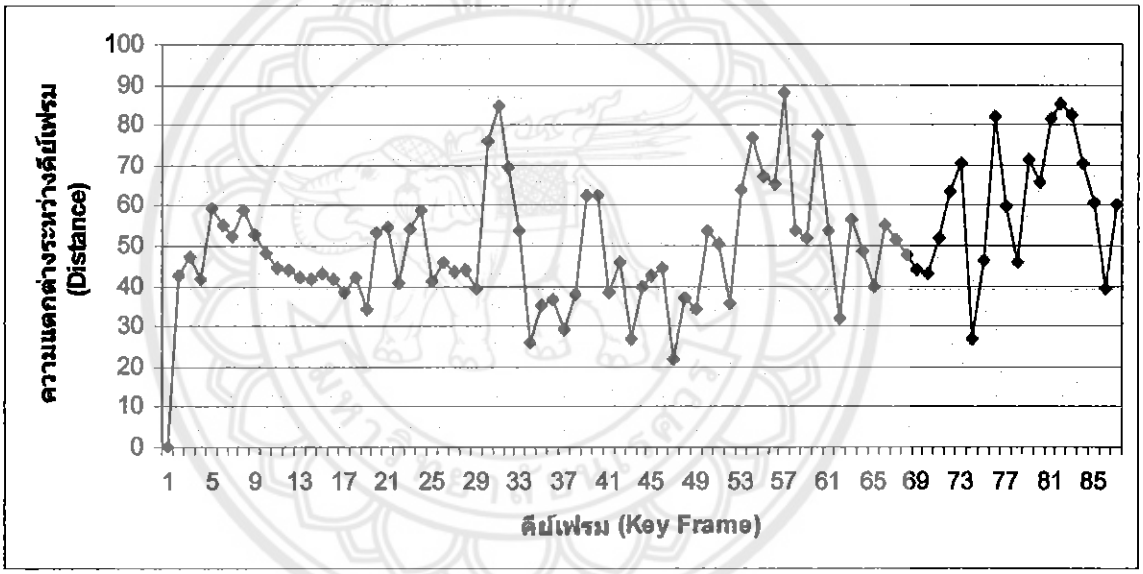
ช่วงเวลาของแต่ละข่าว (วินาทีที่) : 1-50 , 51-220 , 221-306 , 307-368



รูปที่ 4.10 ตัวอย่างภาพในไฟล์วิดีโอข่าว news_ch3_02

ตารางที่ 4.5 ผลลัพธ์การตัดแบ่งวิดีโอในเทรสโฮลด์ระดับต่าง ๆ ของไฟล์วิดีโอ news_ch3_02

ค่าเทรสโฮลด์ ระดับสตอรี่	จำนวนข่าว ที่โปรแกรมตัด	ช่วงเวลาของแต่ละข่าว (วินาทีที่)
15	1	1 -368
17	1	1 -368
20	1	1 -368
25	2	1-200 , 201-368
27	4	1-133 , 134-182 , 183-200 , 201-368
28	5	1-133 , 134-182 , 183-200 , 201-320 , 321-368



รูปที่ 4.11 กราฟความสัมพันธ์ระหว่างคีย์เฟรมของผู้ประกาศข่าวกับคีย์เฟรมที่ตัดในระดับช็อต ของไฟล์วิดีโอ news_ch3_02

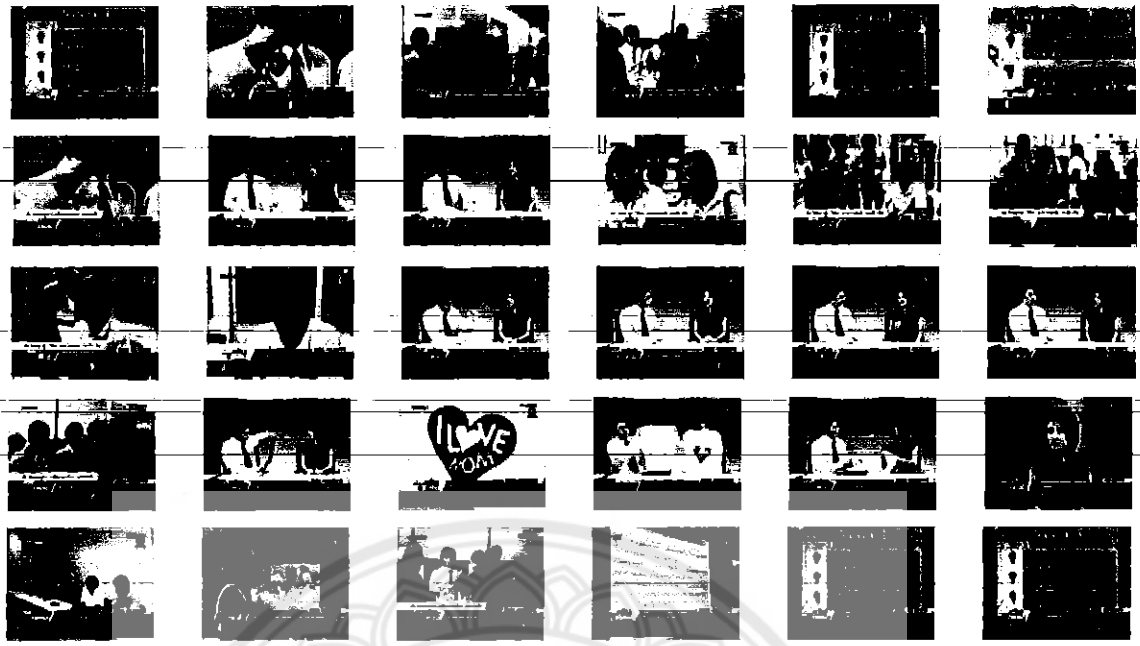
5. ไฟล์วิดีโอข่าว news_ch3_03

ชื่อไฟล์ : news_ch3_03 *ความยาววิดีโอ(วินาที) : 522*

จำนวนเฟรมต่อวินาที : 30 *จำนวนเฟรมทั้งหมด : 15660*

จำนวนช็อต : 71 *จำนวนข่าว : 3*

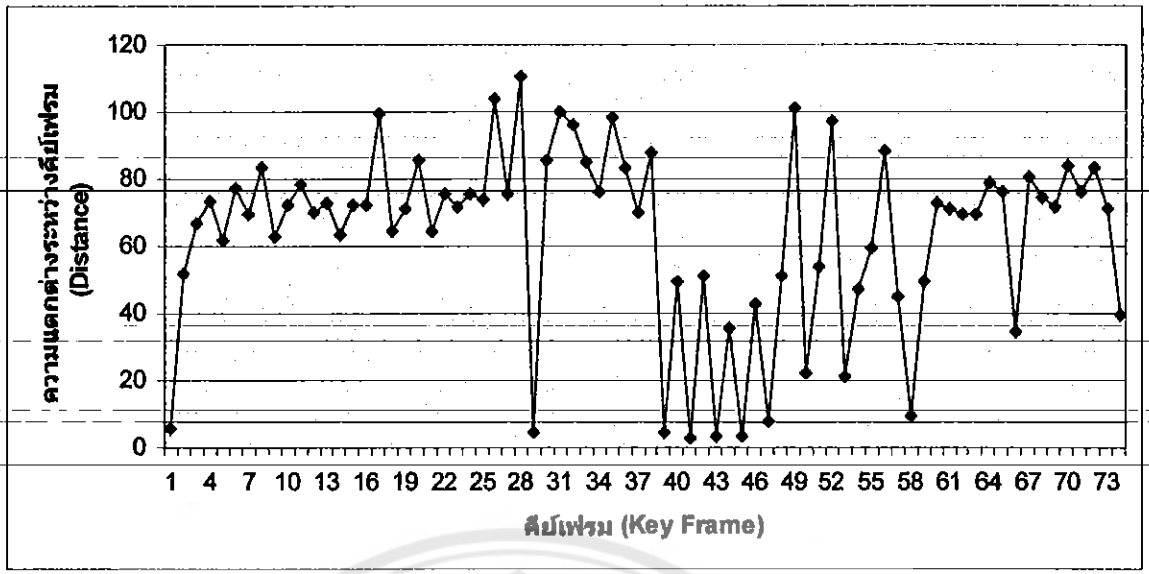
ช่วงเวลาของแต่ละข่าว (วินาทีที่) : 1-205 , 206-312 , 313-522



รูปที่ 4.12 ตัวอย่างภาพในไฟล์วิดีโอข่าว news_ch3_03

ตารางที่ 4.6 ผลลัพธ์การตัดแบ่งวิดีโอในเทรสโฮลด์ระดับต่าง ๆ ของไฟล์วิดีโอ news_ch3_03

ค่าเทรสโฮลด์ ระดับสตอรี่	จำนวนข่าว ที่โปรแกรมตัด	ช่วงเวลาของแต่ละข่าว (วินาทีที่)
3	2	1-357 , 358-522
4	4	1-357 , 358-372 , 373-379 , 380-522
5	6	1-205 , 206-311 , 312-357 , 358-372 , 373-379 , 380-522
10	8	1-205 , 206-311 , 312-357 , 358-372 , 373-379 , 380-395 , 396-438 , 439-522
15	8	1-205 , 206-311 , 312-357 , 358-372 , 373-379 , 380-395 , 396-438 , 439-522
17	8	1-205 , 206-311 , 312-357 , 358-372 , 373-379 , 380-395 , 396-438 , 439-522



รูปที่ 4.13 กราฟความสัมพันธ์ระหว่างคีย์เฟรมของผู้ประกาศข่าวกับคีย์เฟรมที่ตัดในระดับช็อต ของไฟล์วิดีโอ news_ch3_03

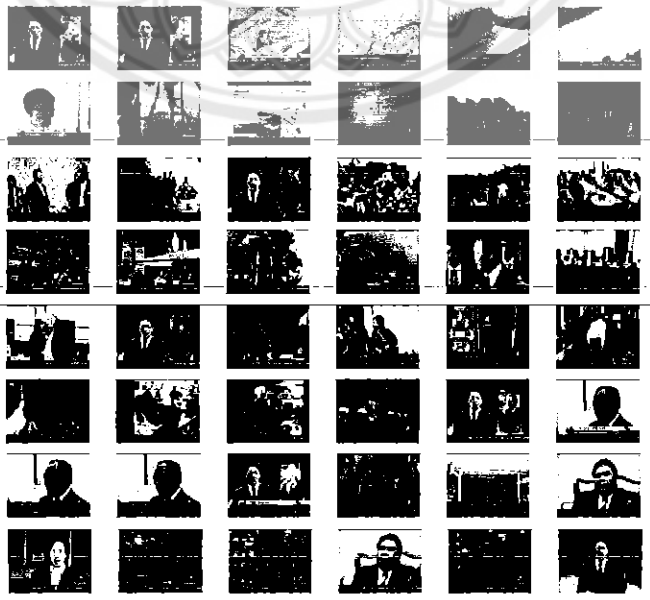
6. ไฟล์วิดีโอข่าว news_modern-9_01

ชื่อไฟล์ : news_modern-9_01 *ความยาววิดีโอ(วินาที) : 1011*

จำนวนเฟรมต่อวินาที : 30 *จำนวนเฟรมทั้งหมด : 30330*

จำนวนช็อต : 22 *จำนวนข่าว : 6*

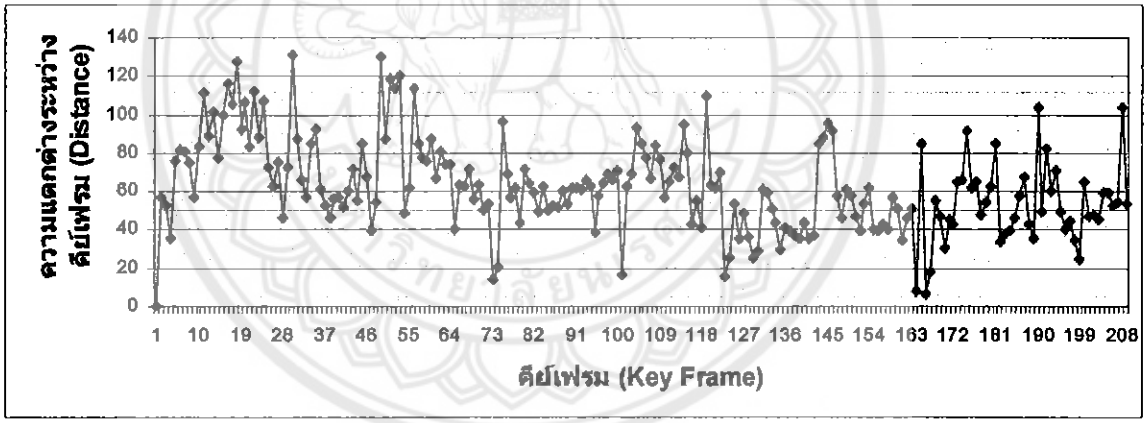
ช่วงเวลาของแต่ละข่าว (วินาทีที่) : 1-320 , 321-427 , 428-560 , 561-733 , 734-818 , 819-1011



รูปที่ 4.14 ตัวอย่างภาพในไฟล์วิดีโอข่าว news_modern-9_01

ตารางที่ 4.7 ผลลัพธ์การตัดแบ่งวิดีโอในเทรสโฮลด์ระดับต่าง ๆ ของไฟล์วิดีโอ news_modern-9_01

ค่าเทรสโฮลด์ ระดับสตอรี่	จำนวนข่าว ที่โปรแกรมตัด	ช่วงเวลาของแต่ละข่าว (วินาทีที่)
5	1	1-1011
10	3	1-734 , 735-814 , 815-1011
13	4	1-320 , 321-734 , 735-814 , 815-1011
17	5	1-320 , 321-557 , 558-734 , 735-814 , 815-1011
18	5	1-320 , 321-557 , 558-734 , 735-814 , 815-1011
19	7	1-320 , 321-427 , 428-557 , 558-734 , 735-814 , 815-816 , 817-1011



รูปที่ 4.15 กราฟความสัมพันธ์ระหว่างคีย์เฟรมของผู้ประกาศข่าวกับคีย์เฟรมที่ตัดในระดับช็อต ของไฟล์วิดีโอ news_modern-9_01

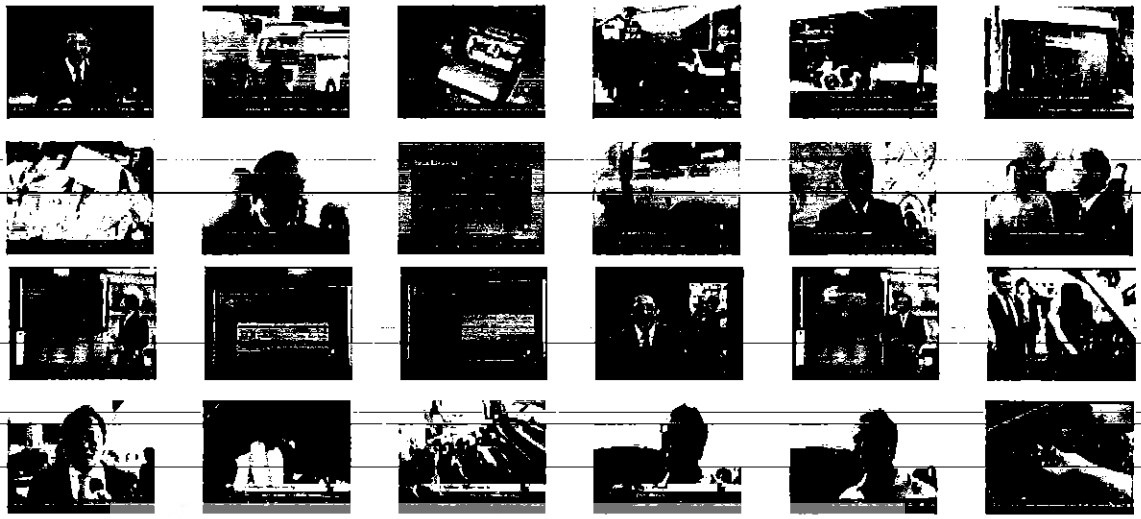
7. ไฟล์วิดีโอข่าว news_modern-9_02 (ใช้ผู้ประกาศคนแรกเป็นคีย์เฟรม)

ชื่อไฟล์ : news_modern-9_02 *ความยาววิดีโอ(วินาที) : 534*

จำนวนเฟรมต่อวินาที : 30 *จำนวนเฟรมทั้งหมด : 16020*

จำนวนช็อต : 86 *จำนวนข่าว : 2*

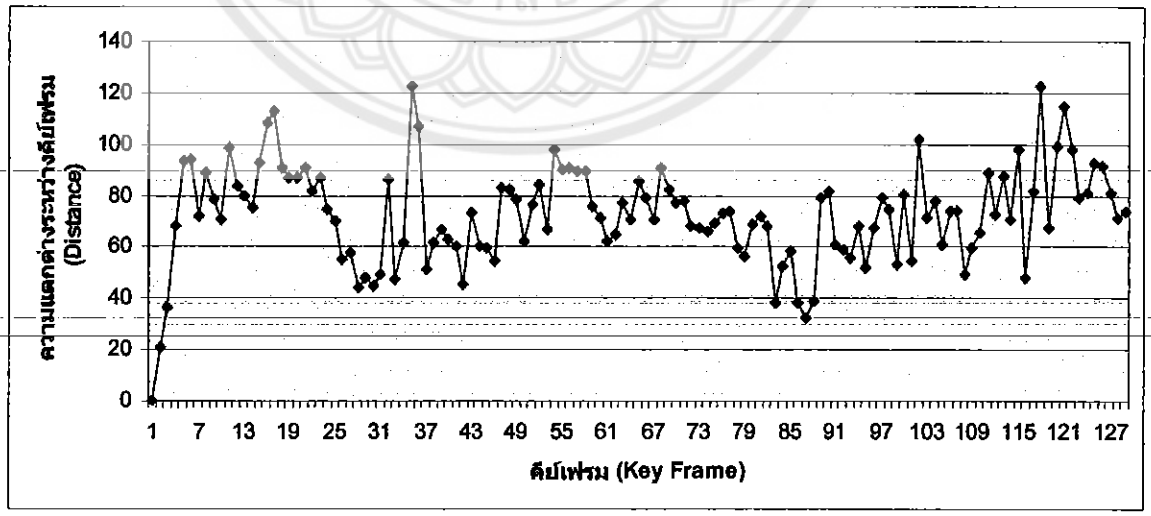
ช่วงเวลาของแต่ละข่าว (วินาทีที่) : 1-369 , 370-534



รูปที่ 4.16 ตัวอย่างภาพในไฟล์วิดีโอข่าว news_modern-9_02

ตารางที่ 4.8 ผลลัพธ์การตัดแบ่งวิดีโอในเทรสโฮลด์ระดับต่าง ๆ ของไฟล์วิดีโอ news_modern-9_02

ค่าเทรสโฮลด์ ระดับสตอรี่	จำนวนข่าว ที่โปรแกรมตัด	ช่วงเวลาของแต่ละข่าว (วินาทีที่)
10	2	1-1 , 2-534
20	2	1-1 , 2-534
25	4	1-1 , 2-369 , 370-371 , 372-534



รูปที่ 4.17 กราฟความสัมพันธ์ระหว่างคีย์เฟรมของผู้ประกาศข่าวคนที่หนึ่งกับคีย์เฟรมที่ตัดในระดับช็อตของไฟล์วิดีโอ news_modern-9_02

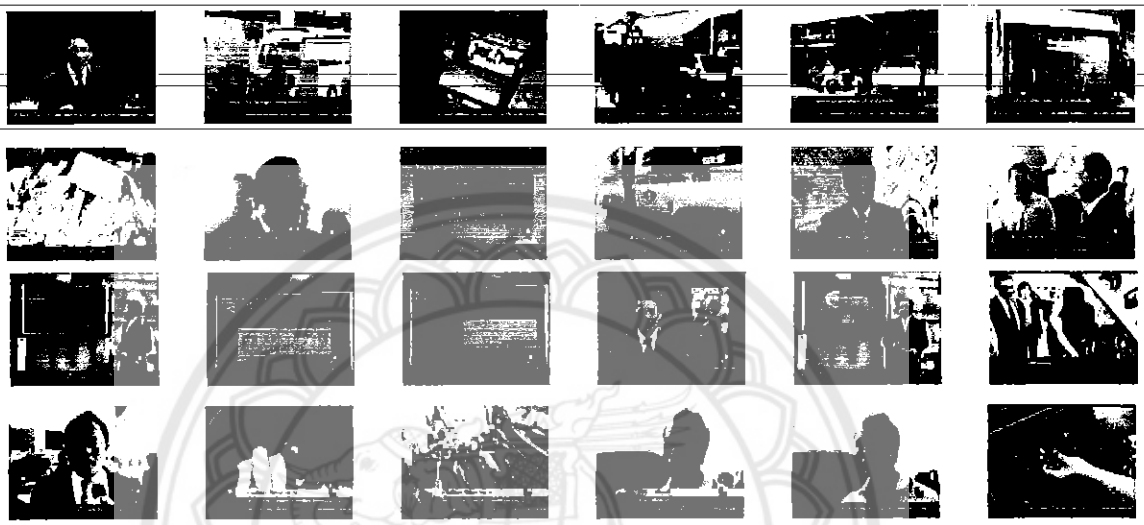
8. ไฟล์วิดีโอข่าว news_modern-9_02 (ใช้ผู้ประกาศคนที่สองเป็นคีย์เฟรม)

ชื่อไฟล์ : news_modern-9_02 *ความยาววิดีโอ(วินาที) : 534*

จำนวนเฟรมต่อวินาที : 30 *จำนวนเฟรมทั้งหมด : 16020*

จำนวนชีอต : 86 *จำนวนข่าว : 5*

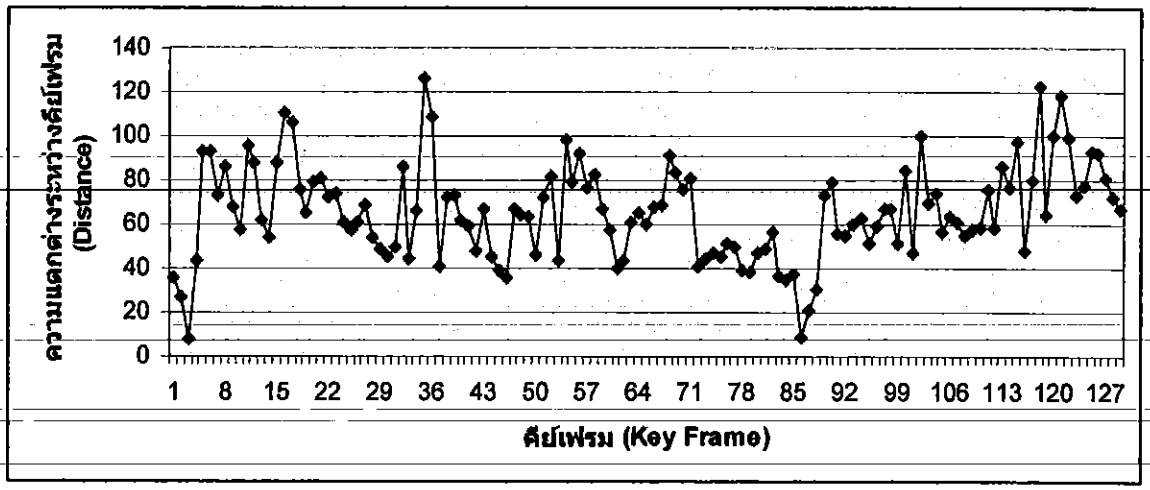
ช่วงเวลาของแต่ละข่าว (วินาทีที่) : 1-27 , 28-207 , 208-304 , 305-369 , 370-534



รูปที่ 4.18 ตัวอย่างภาพในไฟล์วิดีโอข่าว news_modern-9_02

ตารางที่ 4.9 ผลลัพธ์การตัดแบ่งวิดีโอในเทรสโฮลด์ระดับต่าง ๆ ของไฟล์วิดีโอ news_modern-9_02

ค่าเทรสโฮลด์ ระดับสตอรี่	จำนวนข่าว ที่โปรแกรมตัด	ช่วงเวลาของแต่ละข่าว (วินาทีที่)
10	3	1-12 , 13-361 , 362-534
18	3	1-27 , 28-305 , 306-534
20	4	1-27 , 28-207 , 208-305 , 306-534
22	4	1-27 , 28-207 , 208-305 , 306-534
23	4	1-27 , 28-207 , 208-305 , 306-534
25	6	1-27 , 28-207 , 208-305 , 306-323 , 324-332 , 333-534



รูปที่ 4.19 กราฟความสัมพันธ์ระหว่างคีย์เฟรมของผู้ประกาศข่าวคนที่สองกับคีย์เฟรมที่ตัดในระดับช็อตของไฟล์วิดีโอ news_modern-9_02

9. ไฟล์วิดีโอข่าว news_modern-9_03

ชื่อไฟล์ : news_modern-9_03 *ความยาววิดีโอ(วินาที) : 358*

จำนวนเฟรมต่อวินาที : 30 *จำนวนเฟรมทั้งหมด : 10740*

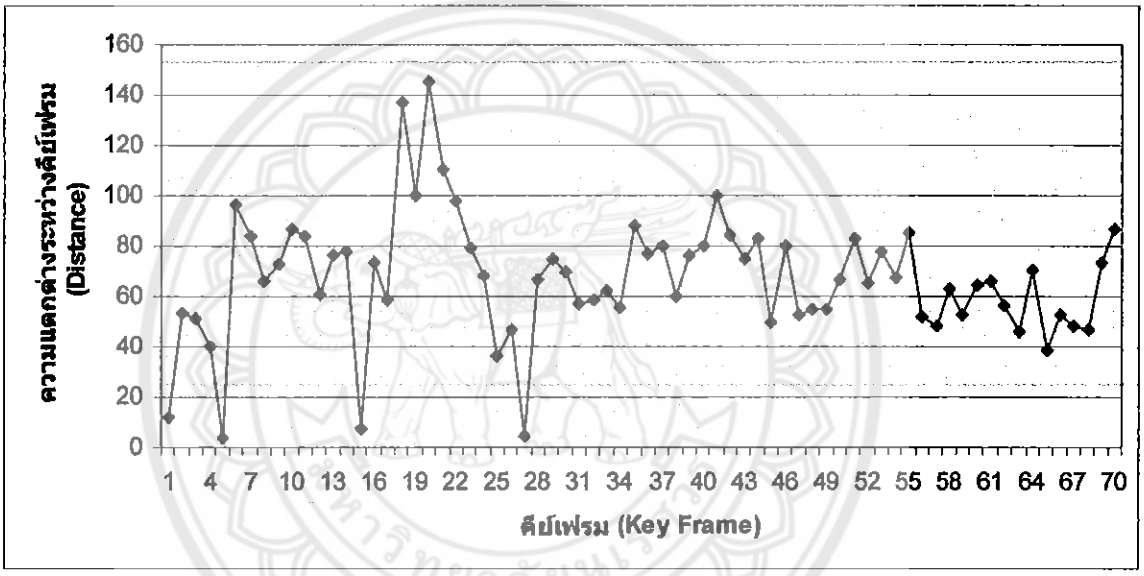
จำนวนช็อต : 73 *จำนวนข่าว : 4*

ช่วงเวลาของแต่ละข่าว (วินาทีที่) : 1-104 , 105-160 , 161-230 , 231-358

รูปที่ 4.20 ตัวอย่างภาพในไฟล์วิดีโอข่าว news_modern-9_03

ตารางที่ 4.10 ผลลัพธ์การตัดแบ่งวิดีโอในเทรสโฮลด์ระดับต่าง ๆ ของไฟล์วิดีโอ news_modern-9_03

ค่าเทรสโฮลด์ ระดับสตอรี่	จำนวนข่าว ที่โปรแกรมตัด	ช่วงเวลาของแต่ละข่าว (วินาทีที่)
5	3	1-105 , 106-231 , 232-358
10	4	1-105 , 106-162 , 163-231 , 232-358
15	4	1-105 , 106-162 , 163-231 , 232-358
25	4	1-105 , 106-162 , 163-231 , 232-358



รูปที่ 4.21 กราฟความสัมพันธ์ระหว่างคีย์เฟรมของผู้ประกาศข่าวกับคีย์เฟรมที่ตัดในระดับช็อต ของไฟล์วิดีโอ news_modern-9_03

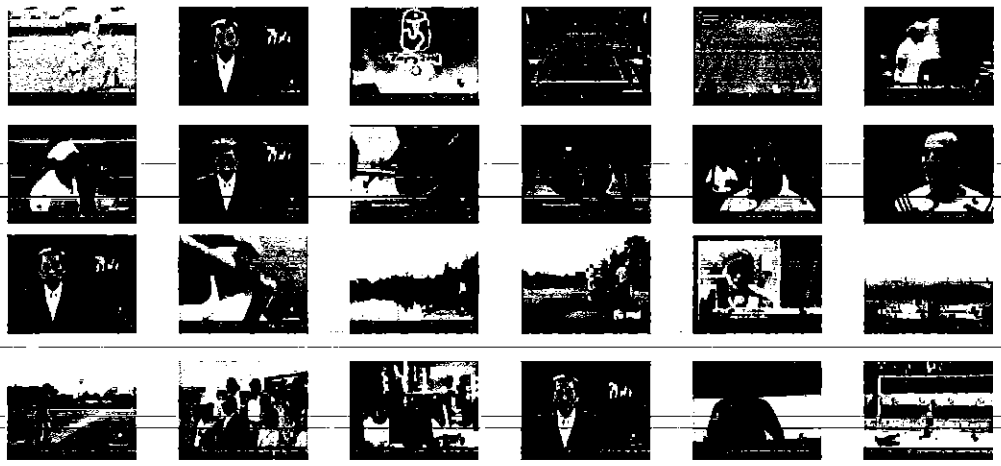
10. ไฟล์วิดีโอข่าว news_modern-9_04

ชื่อไฟล์ : news_modern-9_04 *ความยาววิดีโอ(วินาที) : 428*

จำนวนเฟรมต่อวินาที : 30 *จำนวนเฟรมทั้งหมด : 12840*

จำนวนช็อต : 73 *จำนวนข่าว : 7*

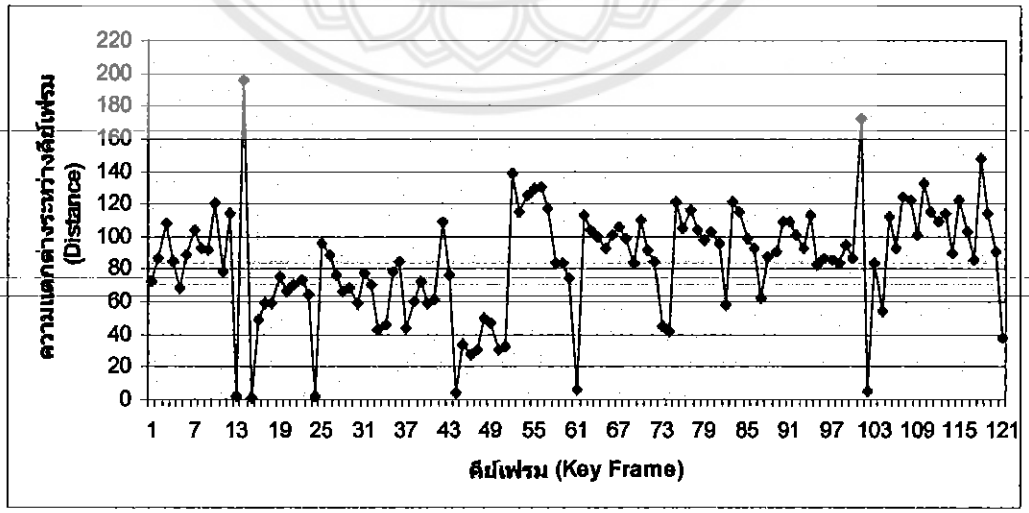
ช่วงเวลาของแต่ละข่าว (วินาทีที่) : 1-33 , 34-57 , 58-109 , 110-158 , 159-220 , 221-370 , 371-428



รูปที่ 4.22 ตัวอย่างภาพในไฟล์วิดีโอข่าว news_modern-9_04

ตารางที่ 4.11 ผลลัพธ์การตัดแบ่งวิดีโอในเทรสโฮลด์ระดับต่าง ๆ ของไฟล์วิดีโอ news_modern-9_04

ค่าเทรสโฮลด์ ระดับสตอรี่	จำนวนข่าว ที่โปรแกรมตัด	ช่วงเวลาของแต่ละข่าว (วินาทีที่)
5	5	1-33 , 34-57 , 58-102 , 103-157 , 158-428
10	7	1-33 , 34-57 , 58-102 , 103-157 , 158-224 , 225-368 , 369-428
15	7	1-33 , 34-57 , 58-102 , 103-157 , 158-224 , 225-368 , 369-428



รูปที่ 4.23 กราฟความสัมพันธ์ระหว่างคีย์เฟรมของผู้ประกาศข่าวกับคีย์เฟรมที่ตัดในระดับช็อตของไฟล์วิดีโอ news_modern-9_04

จากข้อมูลรายละเอียดของการตัดแบ่งวิดีโอระดับสตอรี่ของไฟล์วิดีโอที่ระดับเทรสโฮลด์ต่าง ๆ สามารถนำมาวิเคราะห์และเลือกค่าเทรสโฮลด์มา 3 ค่าเพื่อนำไปทดลอง และหาค่าความถูกต้องของโปรแกรมได้ ดังนี้

ตารางที่ 4.12 ค่าความถูกต้องเมื่อให้ค่าเทรสโฮลด์เท่ากับ 5

ค่าเทรสโฮลด์เท่ากับ 5	จำนวนข่าวของไฟล์วิดีโอ	จำนวนข่าวที่ตัดได้ถูกต้อง	ค่าความถูกต้อง
news_nbt_01	3	1	33.33%
news_nbt_02	2	2	100%
news_ch3_01	3	1	33.33%
news_ch3_02	5	1	20%
news_ch3_03	3	3	100%
news_modern-9_01	6	1	16.66%
news_modern-9_02	4	1	25%
news_modern-9_03	4	3	75%
news_modern-9_04	7	4	66.66%
ไฟล์วิดีโอทั้งหมด	36	17	47.22%

ตารางที่ 4.13 ค่าความถูกต้องเมื่อให้ค่าเทรสโฮลด์เท่ากับ 10

ค่าเทรสโฮลด์เท่ากับ 10	จำนวนข่าวของไฟล์วิดีโอ	จำนวนข่าวที่ตัดได้ถูกต้อง	ค่าความถูกต้อง
news_nbt_01	3	3	100%
news_nbt_02	2	2	100%
news_ch3_01	3	1	33.33%
news_ch3_02	5	1	20%
news_ch3_03	3	3	100%
news_modern-9_01	6	3	50%
news_modern-9_02	4	1	25%
news_modern-9_03	4	4	100%
news_modern-9_04	7	7	100%
ไฟล์วิดีโอทั้งหมด	36	25	69.44%

ตารางที่ 4.14 ค่าความถูกต้องเมื่อให้ค่าเทรสโฮลด์เท่ากับ 15

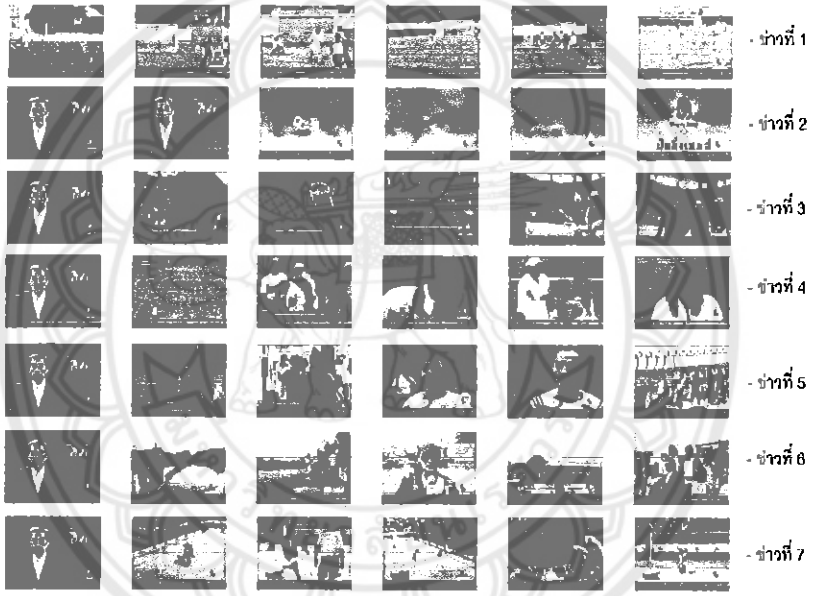
ค่าเทรสโฮลด์เท่ากับ 15	จำนวนข่าวของไฟล์วิดีโอ	จำนวนข่าวที่ตัดได้ถูกต้อง	เปอร์เซ็นต์ความถูกต้อง
news_nbt_01	3	3	100%
news_nbt_02	2	2	100%
news_ch3_01	3	1	33.33%
news_ch3_02	5	1	20%
news_ch3_03	3	3	100%
news_modern-9_01	6	3	50%
news_modern-9_02	4	2	50%
news_modern-9_03	4	4	100%
news_modern-9_04	7	7	100%
ไฟล์วิดีโอทั้งหมด	36	29	80.55%

จากค่าความถูกต้องที่ได้จากตารางที่ 4.12 , 4.13 และ 4.14 จึงตัดสินใจใช้ค่าเทรสโฮลด์เพื่อตัดแบ่งวิดีโอในระดับสตอรี่เป็น 15 เนื่องจากให้ค่าความถูกต้องสูงเมื่อเทียบกับค่าอื่น สามารถตัดวิดีโอได้ถูกต้องมากกว่าค่าอื่นๆที่ได้ทำการทดลอง

ตารางที่ 4.15 เปรียบเทียบผลลัพธ์ของโปรแกรม

ชื่อไฟล์	จำนวนเฟรมที่แตกออกมาได้	วัดด้วยสายตา		โปรแกรม	
		จำนวนช็อต	จำนวนข่าว	จำนวนช็อต	จำนวนข่าว
news_nbt_01	273	51	3	70	3
news_nbt_02	161	32	2	30	2
news_ch3_01	155	22	3	31	1
news_ch3_02	368	74	5	87	1
news_ch3_03	522	71	3	74	3
news_modern-9_01	1011	122	6	209	3
news_modern-9_02	534	86	4	129	2
news_modern-9_03	358	73	4	70	4
news_modern-9_04	428	73	7	121	7

จากตารางที่ 4.15 จะแสดงผลที่ได้จากการวัดด้วยสายตาและผลที่ได้จากการตัดของโปรแกรม สังเกตว่าในการตัดระดับช็อตของโปรแกรมนั้นจะได้จำนวนช็อตมากกว่าค่าที่วัดด้วยสายตา เนื่องจากค่าเทรสโฮลด์ที่เลือกมีค่าต่ำ อาจจะมีเฟรมบางเฟรมที่ไม่ได้มีการเปลี่ยนช็อตแต่โปรแกรมสามารถตัดได้เนื่องจากมีค่าความแตกต่างน้อยกว่า 20 แต่ผลที่ได้จากการตัดในระดับช็อตนี้จะไม่ส่งผล ในการตัดระดับสตอรี่ จึงเลือกใช้ค่าเทรสโฮลด์เท่ากับ 20 เพราะจำทำให้ได้ความละเอียดมาก และในส่วนของการระดับสตอรี่จพบว่าให้โดยทั่วไปให้ค่าความถูกต้องดีกับวิดีโอที่ไม่มีการเปลี่ยนแปลงพื้นหลังหรือผู้ประกาศข่าว แต่สำหรับวิดีโอที่มีการเปลี่ยนพื้นหลังหรอผู้ประกาสข่าวเช่น news_ch3_01 , news_ch3_02 , news_modern-9_01 และ news_modern-9_02 จะได้ผลลัพธ์ที่มีความถูกต้องน้อย



รูปที่ 4.24 ผลลัพธ์ของการตัดแบ่งไฟล์วิดีโอ news_modern-9_04

รูปที่ 4.24 เป็นตัวอย่างผลลัพธ์ที่ได้จากการตัดแบ่งไฟล์วิดีโอของโปรแกรมสังเกตว่า ในส่วนพื้นหลังของผู้ประกาศข่าวจะมีการเปลี่ยนแปลงน้อยมาก ค่าความแตกต่างที่ได้จากการเปรียบเทียบระหว่างเฟรมที่ผู้ใช้เลือกและเฟรมที่ได้จากการตัดในระดับสตอรี่จึงมีค่าน้อย ทำให้ผลที่ได้จากการตัดวิดีโอถูกต้องทั้งหมด

รูปที่ 4.25 ผลลัพธ์ของการตัดแบ่งไฟล์วิดีโอ news_modern-9_01

รูปที่ 4.25 แสดงผลการตัดแบ่งวิดีโอของโปรแกรมของไฟล์ news_modern-9_01 ซึ่งลักษณะของพื้นหลังของผู้ประกาศข่าวมีการเปลี่ยนตามข่าวที่นำเสนอ จะเห็นว่าโปรแกรมจะตัดแบ่งข่าวได้เพียง 3 ข่าวจาก 6 ข่าวเท่านั้น

ตารางที่ 4.16 ค่าความผิดพลาดเมื่อวัดโดย False Positive และ False Negative

ชื่อไฟล์	จำนวนข่าวของไฟล์วิดีโอ	จำนวนข่าวที่ตัดได้	False Positive	False Negative
news_nbt_01	3	4	1/3	-
news_nbt_02	2	3	1/3	-
news_ch3_01	3	1	-	2/3
news_ch3_02	5	1	-	4/5
news_ch3_03	3	8	5/3	-
news_modern-9_01	6	5	-	1/4
news_modern-9_02	4	2	-	2/4
news_modern-9_03	4	4	-	-
news_modern-9_04	7	7	-	-

ตารางที่ 4.16 ค่าความผิดพลาดเมื่อวัดโดย False Positive คือจำนวนวิดีโอที่โปรแกรมตัดได้จำนวนมากกว่าจำนวนข่าวจริง ทั้งนี้เนื่องค่าความแตกต่างระหว่างเฟรมที่ผู้ใช้เลือกกับเฟรมที่ได้จากการตัดระดับชื่อต เฟรมนั้นมีค่าน้อยกว่าค่าเทรสโอลด์ ทำให้โปรแกรมตัดข่าวในเฟรมนั้นด้วย และ False Negative คือจำนวนวิดีโอที่โปรแกรมตัดได้น้อยกว่าจำนวนข่าวจริง เนื่องมาจากค่าความแตกต่างที่ได้จากการเปรียบเทียบระหว่างเฟรมที่ผู้ใช้เลือกกับเฟรมที่ได้จากการตัดระดับชื่อตมีค่า

มากกว่าค่าเทรสโฮลด์ที่กำหนด ค่าความผิดพลาด False Positive และ False Negative จะไม่ได้เกิดจากความผิดพลาดของโปรแกรมแต่จะเกิดจากคุณสมบัติของเฟรมที่นำมาเปรียบเทียบ เช่นลักษณะสีที่ใกล้เคียงหรือต่างกัน หรือแสงแฟลช เป็นต้น

ตารางที่ 4.17 ระยะเวลาการประมวลผลของโปรแกรม

ชื่อไฟล์	ขนาดของวิดีโอ (MB)	ความยาววิดีโอ (วินาที)	ระยะเวลาการประมวลผล (วินาที)
news_nbt_01	11.9	273	129
news_nbt_02	6.98	161	73
news_ch3_01	6.673	155	64
news_ch3_02	15.8	368	179
news_ch3_03	22.4	522	193
news_modern-9_01	45.3	1011	398
news_modern-9_02	23.0	534	220
news_modern-9_03	15.4	358	181
news_modern-9_04	18.4	428	197

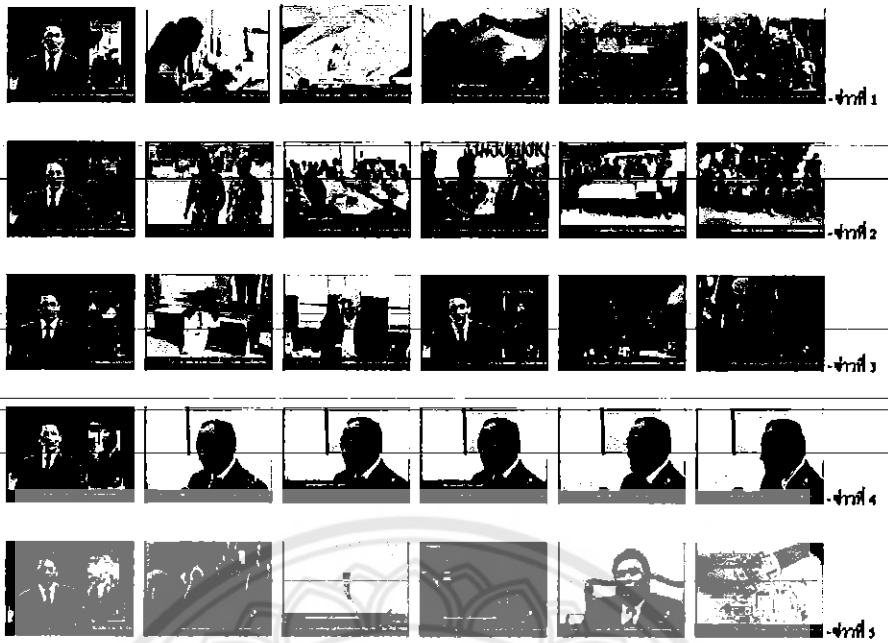
จากตารางที่ 4.17 จะพบว่าระยะเวลาการประมวลผลของโปรแกรมจะแปรผันกับขนาดของไฟล์วิดีโอ หรือความยาวของวิดีโอ เนื่องจากไฟล์ที่มีขนาดใหญ่จะมีจำนวนเฟรมมากทำให้โปรแกรมใช้เวลาในการแตกวิดีโอออกมาเป็นเฟรมมาก และจำนวนเฟรมที่โปรแกรมแตกได้ก็มีมากเช่นกัน ทำให้ระยะเวลาที่ใช้ในการประมวลผลก็จะมากขึ้นตามไปด้วย เช่น news_nbt_02 จะเห็นว่าเป็นไฟล์วิดีโอที่มีขนาดเล็ก จะใช้ระยะเวลาการประมวลผล ส่วน news_modern-9_01 จะมีขนาดไฟล์ใหญ่ทำให้ระยะเวลาการประมวลผลโปรแกรมมาก

ตารางที่ 4.18 ค่าความถูกต้องเมื่อกำหนดให้ผู้ใช้สามารถเลือกเฟรมได้ 3 เฟรมและให้ค่าเทรสโฮลด์เท่ากับ 15

ค่าเทรสโฮลด์เท่ากับ 15	จำนวนข่าวของไฟล์วิดีโอ	จำนวนข่าวที่ตัดได้ถูกต้อง	เปอร์เซ็นต์ความถูกต้อง
news_nbt_01	3	3	100%
news_nbt_02	2	2	100%
news_ch3_01	3	3	100%
news_ch3_02	5	3	60%
news_ch3_03	3	3	100%
news_modern-9_01	6	5	83.33%
news_modern-9_02	4	2	50%
news_modern-9_03	4	4	100%
news_modern-9_04	7	7	100%
ไฟล์วิดีโอทั้งหมด	36	31	86.11%

จากรูปที่ 4.25 จะพบว่าผลลัพธ์จากการตัดวิดีโอจะให้ผลได้ไม่ดีนัก เนื่องจากภาพของผู้ประกาศข่าวมีการเปลี่ยนแปลงตลอด จึงมีการออกแบบให้มีการให้ผู้เลือกเฟรมได้ 3 เฟรมแล้วนำทั้ง 3 เฟรมไปหาความแตกต่างของเฟรมที่ได้จากการตัดในระดับช็อต และเลือกค่าความแตกต่างจากเฟรมที่ให้ค่าน้อยที่สุด ดังที่ได้อธิบายในบทที่ 3

ตารางที่ 4.18 จะแสดงโปรแกรมที่ได้จากการตัดในระดับสตอรี่ โดยให้ผู้ใช้เลือกเฟรมที่จะนำไปเปรียบเทียบในระดับสตอรี่ได้ 3 เฟรม ทั้งนี้เพื่อเพิ่มประสิทธิภาพของโปรแกรม และแก้ปัญหาเรื่องพื้นหลังของผู้ประกาศข่าวมีการเปลี่ยนแปลง รวมถึงมีการเปลี่ยนผู้ประกาศข่าวด้วย เมื่อเปรียบเทียบกับตารางที่ 4.14 จะพบว่าการทำงานของโปรแกรมได้ประสิทธิภาพดีขึ้น เช่นไฟล์วิดีโอ news_modern-9_01 ซึ่งจะมีการเปลี่ยนพื้นหลังตามข่าวที่นำเสนอ จะมีประสิทธิภาพเพิ่มขึ้นอย่างชัดเจน จากที่ตัดได้เพียง 3 ข่าวในการเลือกเฟรมเดียว ก็สามารถตัดได้ 5 ข่าวในการเลือก 3 เฟรม แต่สำหรับ news_ch3_02 และ news_modern-9_02 จะมีการสลับกันเล่าข่าวระหว่างผู้สื่อข่าว 2 คน และมีการสนทนาตอบโต้กัน โดยที่กล้องจะมีการตัดไปส่วนหน้าผู้ประกาศข่าวคนแรกบ้าง คนที่สองบ้าง หรือผู้ประกาศข่าวทั้ง 2 คนบ้าง ซึ่งในส่วนนี้โปรแกรมจะตัดที่เฟรมที่เลือกทั้ง 3 เฟรมได้ถูกต้องแต่บางส่วนจะไม่ใช่ช่วงที่เป็นการนำเสนอข่าว แต่เป็นช่วงที่ผู้ประกาศข่าวมีการสนทนากัน เป็นต้น ทำให้ผลลัพธ์ที่ได้ก็ยังมีความผิดพลาดอยู่บ้าง สำหรับจำนวนช็อตที่ได้จากการตัดในระดับช็อตจะเท่ากับจำนวนช็อตที่ได้ในตารางที่ 4.15



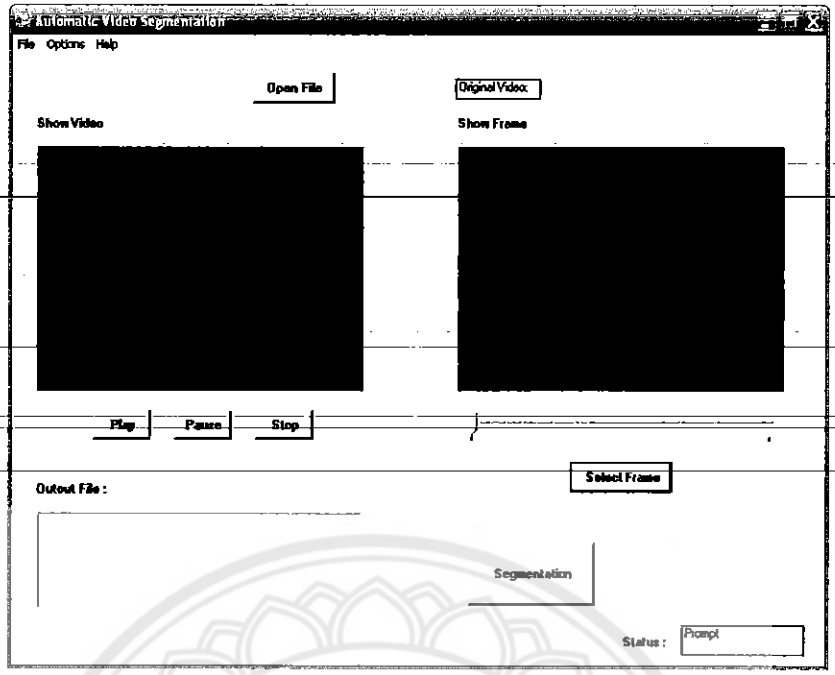
รูปที่ 4.26 ผลลัพธ์ของการตัดแบ่งไฟล์วิดีโอ news_modern-9_01 เมื่อให้ผู้ใช้เลือกได้ 3 เฟรม

จากรูปที่ 4.26 เมื่อเปรียบเทียบกับรูปที่ 4.25 จะพบว่าผลที่ได้จากการตัดวิดีโอของโปรแกรม จะมีประสิทธิภาพดีขึ้น เนื่องจากสามารถตัดวิดีโอได้ถูกต้องมากขึ้น

4.3 การทำงานของโปรแกรม

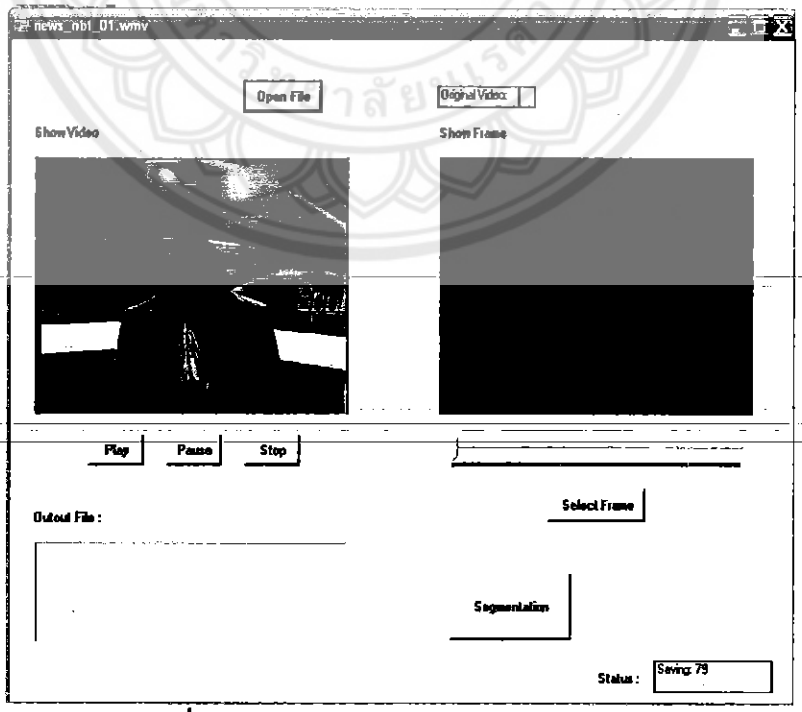
จากหัวข้อ 4.2 ได้ทำการเลือกค่าเทรสโฮลด์ที่เหมาะสมกับกลุ่มไฟล์วิดีโอตัวอย่างแล้ว จะได้โปรแกรมตัดแบ่งวิดีโอข่าวแบบอัตโนมัติ โดยโปรแกรมมีลักษณะการทำงาน ดังนี้

1. เมื่อเปิดโปรแกรมขึ้นมา จะมีลักษณะดังรูปที่ 4.27



รูปที่ 4.27 หน้าตาของโปรแกรม

2. ทำการเปิดไฟล์วิดีโอข่าวที่ต้องการ โดยในที่นี้จะเลือกไฟล์ news_nbt_01 โปรแกรมจะทำการแตกไฟล์วิดีโอเป็นเฟรม (Extract Frame) ให้อัตโนมัติ ดังรูปที่ 4.28 โดยที่แถบ Status ที่มุมขวาด้านล่าง จะทำการอัพเดตว่าขณะนี้โปรแกรมได้ทำการบันทึกไว้ถึงเฟรมที่เท่าใดแล้ว

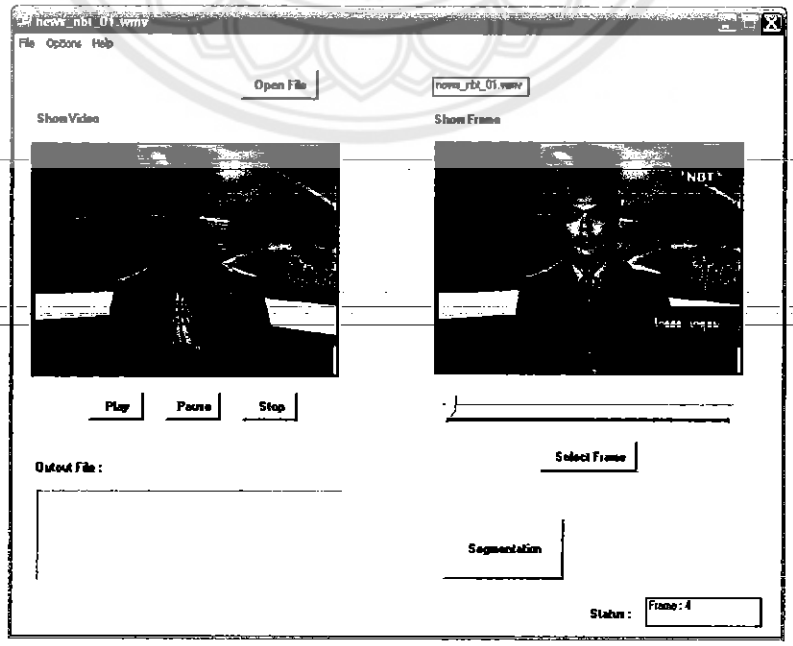


รูปที่ 4.28 โปรแกรมขณะรับไฟล์วิดีโอเข้ามา

3. เมื่อทำการแตกไฟล์วิดีโอเป็นเฟรมเสร็จแล้ว ที่ Status จะแสดงคำว่า Saving DONE

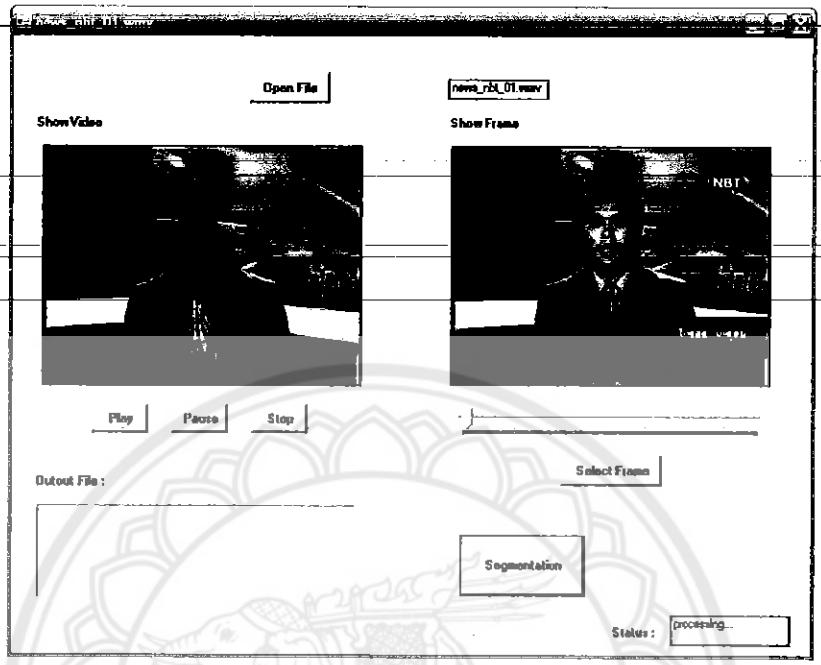
รูปที่ 4.29 โปรแกรมเมื่อทำการแตกไฟล์วิดีโอเป็นเฟรมเสร็จแล้ว

4. เลือกเฟรมของผู้ประกาศข่าวที่ต้องการ เพื่อนำไปเปรียบเทียบในการตัดระดับสตอรี่ โดยทำการเลื่อนที่ Track Bar ที่ Status จะแสดงว่าได้ทำการเลือกเฟรมใด ดังรูปที่ 4.30 เมื่อได้เฟรมที่ต้องการนำไปเปรียบเทียบแล้ว คลิกที่ปุ่ม Select Frame



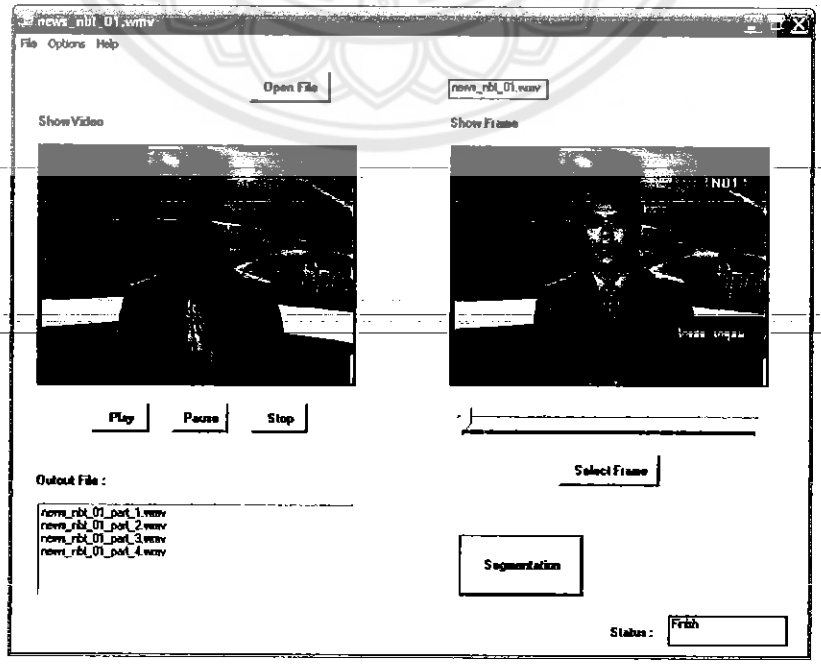
รูปที่ 4.30 โปรแกรมเมื่อทำการเลือกเฟรม

5. คลิกปุ่ม Segmentation เพื่อทำการตัดแบ่งวิดีโอแบบอัตโนมัติ ที่ Status จะแสดงคำว่า processing.. เพื่อแสดงว่ากำลังดำเนินการอยู่ ดังรูปที่ 4.31



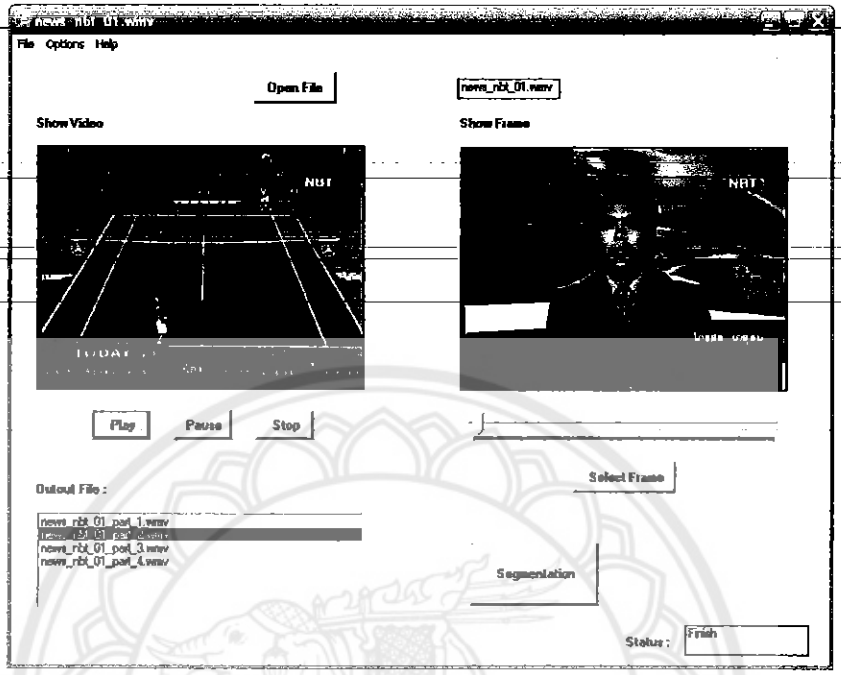
รูปที่ 4.31 โปรแกรมเมื่อทำการกดปุ่ม Segmentation

6. เมื่อโปรแกรมทำการตัดแบ่งวิดีโอข่าวเสร็จแล้ว ที่ Status จะแสดงคำว่า Finish และที่กล่อง Output File จะแสดงไฟล์วิดีโอที่ตัดแบ่งออกมาได้ ดังรูปที่ 4.32



รูปที่ 4.32 โปรแกรมเมื่อดำเนินการเสร็จสิ้น

7. หากต้องการดูไฟล์วิดีโอที่ตัดออกมาได้ ให้คลิกที่ชื่อไฟล์นั้น แล้วคลิกปุ่ม Play ดังรูปที่ 4.33



รูปที่ 4.33 โปรแกรมแสดงวิดีโอเมื่อคลิกเลือก news_nbt_01_part02

บทที่ 5
บทสรุป

5.1 บทสรุป

โครงงานนี้พัฒนาขึ้นเพื่อศึกษาและทดลองเกี่ยวกับอัลกอริทึมในการตัดแบ่งวิดีโอแบบอัตโนมัติ ซึ่งได้มุ่งประเด็นไปที่การนำความสามารถของการจับคู่เปรียบเทียบ (Matching) ของ Color Layout มาใช้หาค่าความแตกต่างของเฟรมภาพที่แตกมาจากไฟล์วิดีโอข่าว โดยทำการเปรียบเทียบ 2 ครั้ง เพื่อให้ความถูกต้องมากที่สุด แล้วทำการตัดแบ่งวิดีโอที่มีเนื้อข่าวมากกว่า 2 ข่าวขึ้นไป ให้เหลือเพียงมีจำนวน 1 ข่าว ต่อ 1 ไฟล์วิดีโอ เพื่อความสะดวกในการจัดเก็บและเปิดดูข่าวย้อนหลัง

โดยโครงงานนี้เป็นโปรแกรมที่พัฒนาขึ้นโดยใช้ Microsoft Visual C# .NET

5.2 สรุปขั้นตอนการทำงาน

รายละเอียดของขั้นตอนการทำงานของโปรแกรม เป็นดังนี้

1. ทำการรับไฟล์วิดีโอข่าวเข้ามา ซึ่งเป็นไฟล์นามสกุล .wmv
2. โปรแกรมจะทำการแตกไฟล์วิดีโอออกเป็นเฟรม แล้วทำการเปรียบเทียบคุณสมบัติของเฟรมภาพที่อยู่ติดกัน โดยใช้การจับคู่เปรียบเทียบ (Matching) ของ Color Layout
3. หากค่าความแตกต่าง (Distance) ของเฟรมภาพคู่ใด อยู่ในช่วงของเทรสโฮลด์ที่กำหนด โปรแกรมจะทำการตัดแบ่งวิดีโอข่าวนั้นในระดับช็อต
4. ทำการเลือกคีย์เฟรมที่เป็นเฟรมของผู้ประกาศข่าวเพื่อใช้เป็นเฟรมหลักในการเปรียบเทียบกับช็อตต่าง ๆ ที่ตัดมาได้
5. เมื่อได้ค่าความแตกต่างมา หากค่าความแตกต่างระหว่างคีย์เฟรมที่เลือกกับเฟรมใด ๆ มีค่าน้อยมาก ๆ แสดงว่าทั้งสองเฟรมนั้นอาจเป็นเฟรมของผู้ประกาศข่าวเหมือนกัน และถ้ามีค่าความแตกต่างต่ำกว่าค่าเทรสโฮลด์ที่กำหนด โปรแกรมจะทำการตัดวิดีโอในระดับสตอรี่ ซึ่งเป็นขั้นตอนสุดท้ายของโปรแกรมที่ได้ให้ผลลัพธ์ของโปรแกรมออกมาเป็นไฟล์วิดีโอย่อย ๆ

ในการทำงานของโปรแกรมยังมีข้อจำกัดอยู่ คือ ไฟล์วิดีโอที่ต่อเนื่องกันไม่ควรยาวมากเกินไปหลายชั่วโมง ทั้งนี้ขึ้นอยู่กับสมรรถนะของเครื่องที่ใช้รันโปรแกรมด้วย

5.3 ปัญหาและอุปสรรคในการพัฒนา

5.3.1 ผลของการทำงานโปรแกรมแปรผันกับการเปลี่ยนแปลงการเข้ารหัส (Encode) ข้อมูล ส่งผลให้การเปลี่ยนแปลงการเข้ารหัสข้อมูลแต่ละครั้งอาจทำให้ผลที่ได้ต่างกัน

5.3.2 องค์ประกอบและสิ่งแวดล้อมของภาพส่งผลต่อคุณสมบัติของวิดีโอและเฟรมภาพที่ได้ เช่น ความสว่าง , แสงแฟลชในวิดีโอทำให้ค่า Color Layout มีการเปลี่ยนแปลง ทำให้ผลของโปรแกรมเปลี่ยน

5.3.3 วิดีโอแต่ละไฟล์มีคุณสมบัติต่างกันออกไป ทำให้การเลือกค่าเทรสโฮลด์ ต้องเลือกค่าที่เหมาะสมมากที่สุด

5.3.4 หากพื้นหลังของผู้ประกาศข่าวมีความเปลี่ยนแปลงมากก็จะทำให้การเปรียบเทียบเกิดความคลาดเคลื่อนมาก ส่งผลให้ทำการตัดแบ่งวิดีโอเกิดความผิดพลาด

5.4 ข้อเสนอแนะ

เนื่องจากโครงงานนี้ เป็นการเริ่มพัฒนาในขั้นแรก ทำให้อาจจะยังมีข้อผิดพลาดของโปรแกรมอยู่บ้าง หากผู้ที่สนใจพัฒนาโครงงานในขั้นต่อ ๆ ไป อาจมีข้อเสนอแนะดังนี้

5.4.1 การเพิ่มทฤษฎีเกี่ยวกับภาพและวิดีโอแบบต่าง ๆ เพื่อใช้ในการดึงคุณลักษณะต่าง ๆ ของวิดีโอนอกจากสีมาทำการเปรียบเทียบด้วย

5.4.2 มีการใช้เสียงในวิดีโอข่าวมาทำการประมวลผลด้วย จะทำให้มีความแม่นยำมากขึ้น

5.4.3 การเพิ่มสรรถนะของโปรแกรมให้รองรับไฟล์วิดีโอที่มีความยาวมาก ๆ ได้ แต่ทั้งนี้ขึ้นกับเครื่องคอมพิวเตอร์ที่ใช้รันโปรแกรมด้วย

5.5 แนวทางในการพัฒนาต่อในอนาคต

5.5.1 สามารถประยุกต์ใช้กับวิดีโอรูปแบบอื่น ๆ นอกจากวิดีโอข่าวได้

5.5.2 สามารถที่จะเพิ่มวิธีการในการเปรียบเทียบภาพได้มากขึ้น โดยอาศัยคุณสมบัติอื่น ๆ ของ MPEG-7

5.5.3 หากต้องการเพิ่มความถูกต้องของโปรแกรม สามารถทดลองซ้ำหลาย ๆ ครั้งเพื่อหาค่าเทรสโฮลด์ที่เหมาะสมได้ หรือเพิ่มวิธีการในการเปรียบเทียบเพื่อให้มีความแม่นยำมากยิ่งขึ้น

5.6 สรุปผลการดำเนินการ

การพัฒนาโปรแกรมในการตัดแบ่งวิดีโอข่าวแบบอัตโนมัติ สามารถสรุปได้ว่าการตัดแบ่งวิดีโอสามารถทำได้ดี แต่มีขอบเขตจำกัดดังที่กล่าวไปแล้วข้างต้น และผลของโปรแกรมจะขึ้นอยู่กับความละเอียดในการแปลงข้อมูลไฟล์วิดีโอ และคุณสมบัติของภาพ

เอกสารอ้างอิง

[1] " ความเป็นมาของมัลติมีเดีย " [Online]. Available :
http://www.thaigoodview.com/library/contest2551/tech04/21/standard/m01.html 2008.

[2] " วิดีโอ (Video) " [Online]. Available :
www.computer.cmru.ac.th/~pornwana/UserFiles/File/multimedia_doc/unit9.pdf

[3] Sarah Porter. " Video Segmentation and Indexing using Motion Estimation. ". PhD thesis.
University of Bristol. February 2004.

[4] " การผลิตรายการโทรทัศน์เบื้องต้น " [Online]. Available :
http://regelearning.payap.ac.th/docu/ca425/chapter/chapter4_0.htm 2008.

[5] " YCbCr Definition " [Online]. Available :
http://www.pcmag.com/encyclopedia_term/0,2542,t=YCbCr&i=55147,00.asp

[6] Sonera MediaLab. " MPEG-7 White Paper. " [Online]. Available :
http://www.medialab.sonera.fi/workspace/MPEG7WhitePaper.pdf 2008.

[7] Neli Day & Jose M.Martinez. " Introduction to MPEG-7 (v4.0) " [Online]. Available :
http://xml.coverpages.org/MPEG7-N4675-Intro.pdf 2008.

ประวัติผู้เขียนโครงงาน

ชื่อ นางสาวนันทวดี โฉมยงค์

ภูมิลำเนา 53/2 ม.11 ต.ทุ่งเสลี่ยม อ.ทุ่งเสลี่ยม จ.สุโขทัย 64150

ประวัติการศึกษา

- จบระดับมัธยมศึกษาจากโรงเรียนทุ่งเสลี่ยมชนูปถัมภ์
- ปัจจุบันกำลังศึกษาในระดับปริญญาตรีชั้นปีที่ 4

สาขาวิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์

มหาวิทยาลัยนเรศวร

E-mail : beau_9198@hotmail.com

ชื่อ นางสาวปวีณา สารชาติ

ภูมิลำเนา 58/7 ม.12 ต.สระกรวด อ.ศรีเทพ จ.เพชรบูรณ์ 67170

ประวัติการศึกษา

- จบระดับมัธยมศึกษาจากโรงเรียนศรีเทพประชาสรรค์
- ปัจจุบันกำลังศึกษาในระดับปริญญาตรีชั้นปีที่ 4

สาขาวิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์

มหาวิทยาลัยนเรศวร

E-mail : paweenah@hotmail.com