

เทคนิคการควบคุมอัตราสำหรับการเข้ารหัสวีดิทัศน์สตรีมมิงแบบออนไลน์
บนพื้นฐานของมาตรฐาน H.264



นายสิริโรจน์ ก้อนในเมือง

สถาบันวิทยบริการ

จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมไฟฟ้า ภาควิชาวิศวกรรมไฟฟ้า

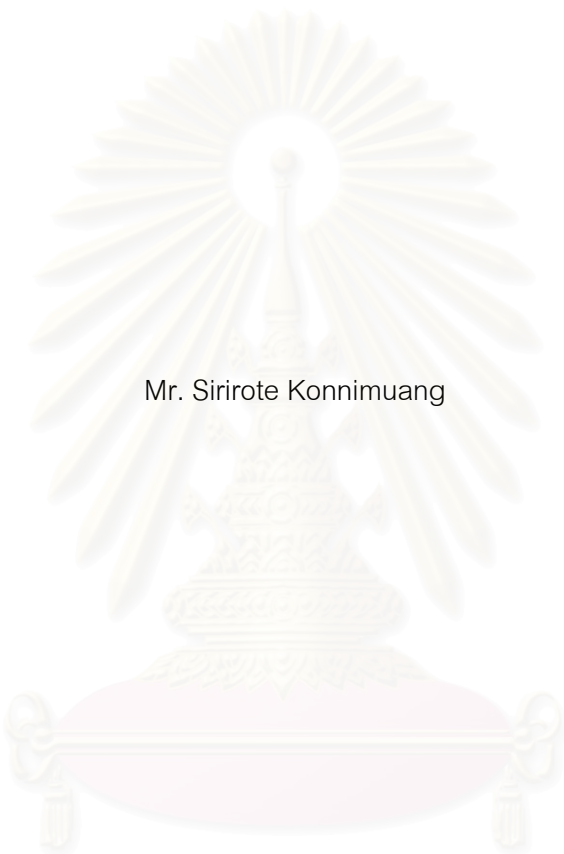
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2547

ISBN 974-53-1748-9

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

RATE CONTROL TECHNIQUE FOR OFFLINE STREAMING VIDEO ENCODING
BASED ON H.264 STANDARD



Mr. Sirote Konnimuang

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Engineering in Electrical Engineering

Department of Electrical Engineering

Faculty of Engineering

Chulalongkorn University

Academic Year 2004

ISBN 974-53-1748-9

สิริโรจน์ ก้อนในเมือง : เทคนิคการควบคุมอัตราสำหรับการเข้ารหัสวิดีโอที่สนับสนุนการเข้ารหัสแบบออฟไลน์บนพื้นฐานของมาตรฐาน H.264 (RATE CONTROL TECHNIQUE FOR OFFLINE STREAMING VIDEO ENCODING BASED ON H.264 STANDARD)

อ. ที่ปรึกษา : ผู้ช่วยศาสตราจารย์ ดร. สุภาวดี อร่ามวิทย์, 102 หน้า. ISBN 974-53-1748-9.

วิทยานิพนธ์นี้นำเสนอเทคนิคการควบคุมอัตราสำหรับการเข้ารหัสวิดีโอที่สนับสนุนการเข้ารหัสแบบออฟไลน์บนพื้นฐานของมาตรฐาน H.264 เพื่อให้คุณภาพของสัญญาณวิดีโอที่ขึ้น เทคนิคที่นำเสนอนี้จะประยุกต์ใช้ระเบียบวิธีการเข้ารหัสแบบหน้าต่างเลื่อน (Sliding Window Encoding Scheme) ซึ่งจะใช้ข้อมูลทางสถิติของเฟรมในอนาคตมาช่วยในการจัดสรรบิต และวิธีการประมาณบิตเป้าหมายแบบปรับตัวได้ตามความซับซ้อนของการเข้ารหัสในแต่ละเฟรม โดยการใช้อัตราส่วนค่าความแตกต่างสัมบูรณ์เฉลี่ย (MAD ratio) ผลการทดลองแสดงให้เห็นว่าตัวเข้ารหัส H.264 ที่ใช้ระเบียบวิธีที่นำเสนอสามารถลดการขึ้น ๆ ลง ๆ และตกลงอย่างฉับพลันของค่า PSNR สำหรับเฟรมที่มีการเคลื่อนไหวมากหรือมีการเปลี่ยนแปลง ในการเข้ารหัสวิดีโอที่สนับสนุนการ 2 ข้อที่ต้องถูกพิจารณา กล่าวคือ เวลาบรรจุก่อนที่ผู้รับชมวิดีโอต้องรอ และขนาดของบัฟเฟอร์ของตัวถอดรหัสที่ภาครับ โดยการพิจารณา 2 ความต้องการ วิดีโอที่เข้ารหัสยังคงมีคุณภาพวิดีโอที่ดีกว่าเมื่อเปรียบเทียบกับวิดีโอที่เข้ารหัสด้วย H.264

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชา.....วิศวกรรมไฟฟ้า.....ลายมือชื่อ.....

สาขาวิชา.....วิศวกรรมไฟฟ้า.....ลายมือชื่ออาจารย์ที่ปรึกษา.....

ปีการศึกษา.....2547.....ลายมือชื่ออาจารย์ที่ปรึกษาร่วม.....

4570596821: MAJOR ELECTRICAL ENGINEERING

KEY WORD: BUFFER / OFF-LINE / RATE CONTROL / STREAMING VIDEO / H.264

SIRIROTE KONNIMUANG: RATE CONTROL TECHNIQUE FOR OFFLINE
STREAMING VIDEO ENCODING BASED ON H.264 STANDARD. THESIS

ADVISOR: ASST. PROF. SUPAVADEE ARAMVITH, Ph.D., 102 pp.

ISBN 974-53-1478-9.

The thesis proposed rate control technique for offline streaming video encoding based on H.264 standard to enhance the quality of video signal. The technique proposed applies a sliding window rate-control scheme which utilizes statistical information of the future video frames with an adaptive estimation of target bits according to the coding complexities of each frame by using mean absolute difference (MAD). Simulation results show that the H.264 encoder integrated with our proposed algorithm effectively alleviates peak signal to noise ratio (PSNR) surges and sharp drops a frame that has high motion or scene change. During the encoding process for streaming video, two requirements need to be considered: the pre-loading time that the video viewers have to wait and the physical buffer-size at the receiver side. By considering both requirements, the video encoded still has better video quality compared to H.264.



Department.....Electrical Engineering..... Student's Signature.....

Field of study.....Electrical Engineering..... Advisor's signature.....

Academic year.....2004.....Co-advisor's signature.....

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยความช่วยเหลืออย่างดียิ่งของท่านอาจารย์ที่ปรึกษา วิทยานิพนธ์ ผู้ช่วยศาสตราจารย์ ดร. สุภาวดี อร่ามวิทย์ ซึ่งได้ให้ความกรุณาแนะนำและ เสนอแนะ ข้อคิดเห็นต่าง ๆ ในการทำวิจัยมาตลอด ผู้วิจัยจึงขอกราบขอบพระคุณ

กราบขอบพระคุณบิดามารดา ผู้ซึ่งทำให้ผู้วิจัยได้รับแต่สิ่งดี ๆ เสมอมา ขอบพระคุณ สำหรับความรัก ความห่วงใยอาทรและกำลังใจที่ไม่เคยลดน้อยลงไปเลย

ขอขอบคุณน้อง ๆ ในห้องปฏิบัติการวิจัยทุก ๆ คน ที่ให้คำปรึกษาในด้านการเขียน โปรแกรม รวมถึงเพื่อน ๆ และคนรอบตัวของผู้วิจัยทุก ๆ กลุ่ม และทุก ๆ คน ไม่ว่าจะเป็นที่อยู่ ภายในห้องปฏิบัติการวิจัยกรรมวิธีสัญญาณดิจิทัล (Digital Signal Processing Research Laboratory) หรือที่ได้ก็ตาม สำหรับความช่วยเหลือ และแรงบันดาลใจ ขอขอบคุณที่ทำให้วันเวลา ของผู้วิจัยผ่านไปอย่างมีความหมาย

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญตาราง.....	ญ
สารบัญภาพ.....	ฎ
บัญชีคำศัพท์.....	ฒ
บทที่	
1. บทนำ.....	1
1.1 แนวเหตุผลและความเป็นมา.....	1
1.2 มาตรฐานการเข้ารหัสสัญญาณวีดิทัศน์.....	4
1.2.1 มาตรฐานของ ITU-T.....	4
1.2.2 มาตรฐานของ ISO/IEC.....	5
1.3 แนวทางที่นำเสนอ.....	7
1.4 วัตถุประสงค์.....	7
1.5 ขอบเขตของวิทยานิพนธ์.....	8
1.6 แนวทางการดำเนินงาน.....	8
1.7 ประโยชน์ที่คาดว่าจะได้รับ.....	8
2 ความรู้พื้นฐาน.....	9
2.1 วีดิทัศน์ดิจิทัล.....	9
2.2 การเข้ารหัสสัญญาณวีดิทัศน์.....	10
2.2.1 ชนิดของรูปภาพ.....	11
2.2.2 การเข้ารหัสภายในเฟรม.....	12
2.2.3 การเข้ารหัสระหว่างเฟรม.....	16
2.2.4 การแปลงและอินเวอร์สการแปลง.....	17
2.2.5 การควอนไทซ์และการสแกนค่าสัมประสิทธิ์ที่ได้.....	18
2.2.6 การเข้ารหัสเอนโทรปี.....	19

2.2.7	การประมาณการเคลื่อนที่ และการชดเชยการเคลื่อนที่.....	19
2.2.8	รูปแบบการจัดเรียงสัญญาณ.....	23
2.2.9	รูปแบบสัญญาณวีดิทัศน์ที่ใช้.....	24
2.3	การควบคุมอัตราสำหรับการเข้ารหัสสัญญาณวีดิทัศน์ H.264.....	25
2.3.1	การควบคุมอัตราระดับกลุ่มภาพ.....	25
2.3.2	การควบคุมอัตราระดับเฟรม.....	28
2.3.2.1	ส่วน Pre-Encoding Stage	28
2.3.2.2	ส่วน Post-Encoding Stage	32
2.3.3	การควบคุมอัตราระดับหน่วยพื้นฐานย่อย.....	32
2.3.3.1	นิยามของหน่วยพื้นฐานย่อย.....	32
2.3.3.2	การควบคุมอัตราของหน่วยพื้นฐานย่อย.....	33
2.4	การวัดคุณภาพของภาพ.....	35
2.4.1	ค่าความผิดพลาดกำลังสองเฉลี่ย.....	36
2.4.2	ค่าอัตราส่วนสัญญาณยอดต่อสัญญาณรบกวน.....	36
2.5	การทบทวนงานวิจัยที่เกี่ยวข้อง.....	37
2.5.1	ระเบียบวิธีการเข้ารหัสแบบหน้าต่างเลื่อน.....	37
2.5.2	ความจำเป็นของขนาดบัฟเฟอร์ที่ตัวถอดรหัส และเวลาบรรจุก่อน.....	38
2.5.2.1	การหาขนาดของบัฟเฟอร์ตัวถอดรหัสและเวลาบรรจุก่อนจาก บิตสตรีมวีดิทัศน์.....	40
2.5.2.2	การผลิตบิตสตรีมวีดิทัศน์ที่ได้เข้ารหัสแล้วเมื่อมีการกำหนด ขนาดของ บัฟเฟอร์ตัวถอดรหัสและเวลาบรรจุก่อน.....	41
2.5.3	การประมาณบิตเป้าหมายแบบปรับตัวได้.....	42
3	รายละเอียดเกี่ยวกับวิทยานิพนธ์และระบบที่นำเสนอ.....	43
3.1	ภาพรวมของระบบ.....	43
3.2	ลักษณะสัญญาณวีดิทัศน์ขาเข้า.....	44
3.3	การทำงานของระบบที่นำเสนอ.....	45
3.3.1	ระเบียบวิธีการควบคุมอัตราที่นำเสนอ.....	45
3.3.1.1	การเข้ารหัสครั้งแรก.....	46

3.3.1.2 การเข้ารหัสครั้งที่สอง.....	46
3.3.1.3 ระเบียบวิธีการควบคุมอัตราที่นำเสนอกับข้อกำหนดของขนาด บัพเฟอร์ตัวถอดรหัสและเวลาบรจุก่อน.....	48
4 ผลการทดสอบ.....	52
4.1 ขั้นตอนการทดสอบ.....	52
4.1.1 การกำหนดพารามิเตอร์ในการเข้ารหัส.....	52
4.1.2 ลำดับวิธีทัศน์ที่ใช้ในการทดสอบ.....	53
4.2 ผลการทดสอบและวิเคราะห์ผลการทดสอบ.....	54
4.2.1 การเข้ารหัสครั้งแรก.....	54
4.2.2 ผลการทดสอบเมื่อไม่มีข้อกำหนดของขนาดบัพเฟอร์ตัวถอดรหัส และเวลาบรจุก่อน.....	56
4.2.1.1 ผลการทดสอบกับลำดับสัญญาณวิธีทัศน์ Foreman	56
4.2.1.2 ผลการทดสอบกับลำดับสัญญาณวิธีทัศน์ Carphone	68
4.2.1.3 ผลการทดสอบกับลำดับสัญญาณวิธีทัศน์ Suzie+Trevor	80
4.2.3 ผลการทดสอบเมื่อมีข้อกำหนดของขนาดบัพเฟอร์ตัวถอดรหัส และเวลาบรจุก่อน.....	92
5 สรุปผลการวิจัยและข้อเสนอแนะ.....	96
5.2 สรุปผลการวิจัย.....	96
5.3 ข้อเสนอแนะสำหรับการวิจัยในอนาคต.....	96
รายการอ้างอิง.....	98
ประวัติผู้เขียนวิทยานิพนธ์.....	102

สารบัญตาราง

	หน้า
ตารางที่ 1.1 มาตรฐานการเข้ารหัสสัญญาณวีดิทัศน์ของ ITU-T.....	5
ตารางที่ 1.2 มาตรฐานการเข้ารหัสสัญญาณวีดิทัศน์ของ ISO/IEC.....	5
ตารางที่ 2.1 รูปแบบของสัญญาณวีดิทัศน์ที่ใช้งานในปัจจุบัน.....	9
ตารางที่ 4.1 พารามิเตอร์ที่ใช้ในการทดสอบ.....	53
ตารางที่ 4.2 ผลการเปรียบเทียบคุณภาพของลำดับวีดิทัศน์ Foreman สำหรับการเข้ารหัส JM6.1 และระเบียบวิธีที่นำเสนอ ที่ขนาดหน้าต่างต่าง $W = 50$ และ 100	66
ตารางที่ 4.3 ผลการเปรียบเทียบคุณภาพของลำดับวีดิทัศน์ Carphone สำหรับการเข้ารหัส JM6.1 และระเบียบวิธีที่นำเสนอ ที่ขนาดหน้าต่างต่าง $W = 50$ และ 100	78
ตารางที่ 4.4 ผลการเปรียบเทียบคุณภาพของลำดับวีดิทัศน์ Suzie+Trevor สำหรับการเข้ารหัส JM6.1 และระเบียบวิธีที่นำเสนอ ที่ขนาดหน้าต่างต่าง $W = 50$ และ 100	90
ตารางที่ 4.5 ผลการเปรียบเทียบคุณภาพของลำดับวีดิทัศน์ Suzie+Trevor สำหรับการเข้ารหัส JM6.1 และระเบียบวิธีที่นำเสนอ ที่ขนาดหน้าต่างต่าง $W = 50$ และ 100 เมื่อมีข้อกำหนดของเวลาบรจูก่อน และขนาดของบัพเฟอรัตัวถอดรหัส.....	94

สารบัญภาพ

หน้า

รูปที่ 1.1 ผู้ใช้สามารถเข้าถึงฐานข้อมูลวีดิทัศน์บนเครือข่ายสำหรับงาน ประยุกต์วีดิทัศน์สตรีมมิง.....	3
รูปที่ 2.1 บล็อกไดอะแกรมของการเข้ารหัสสัญญาณวีดิทัศน์ที่มีการบีบอัด.....	10
รูปที่ 2.2 ความสัมพันธ์ระหว่างเฟรมแบบ I P และ B ในการทำนายการเคลื่อนที่.....	12
รูปที่ 2.3 ส่วนประกอบความส่องสว่างขนาด 4 x 4 จุดภาพ.....	13
รูปที่ 2.4 ตำแหน่งส่วนประกอบความส่องสว่างขนาด 4 x 4 จุดภาพ.....	14
รูปที่ 2.5 การเข้ารหัสภายในเฟรมของส่วนประกอบความส่องสว่างขนาด 4x4 จุดภาพใหม่ 0 ถึง 8	15
รูปที่ 2.6 เข้ารหัสภายในเฟรมของส่วนประกอบความส่องสว่างขนาด 16x16 จุดภาพใหม่ 0 ถึง 3	16
รูปที่ 2.7 การจัดเรียงค่าสัมประสิทธิ์.....	17
รูปที่ 2.8 วิธีการรูดค่าสัมประสิทธิ์แบบซิกแซกและแบบคู่.....	18
รูปที่ 2.9 การแบ่งบล็อกขนาด 16x16 จุดภาพ.....	19
รูปที่ 2.10 การแบ่งบล็อกขนาด 8x8 จุดภาพ.....	20
รูปที่ 2.11 แสดงการแบ่งบล็อกภายในภาพ.....	20
รูปที่ 2.12 การประมาณการเคลื่อนที่.....	21
รูปที่ 2.13 การหาค่าจุดภาพระหว่าง 2 จุดภาพ.....	22
รูปที่ 2.14 การจัดกลุ่มมาโครบล็อกแบบยืดหยุ่น.....	23
รูปที่ 2.15 การเรียงลำดับมาโครบล็อกแบบปกติ.....	24
รูปที่ 2.16 บล็อกไดอะแกรมการควบคุมอัตราสำหรับมาตรฐาน H.264	27
รูปที่ 2.17 ขนาดหน่วยพื้นฐานย่อย.....	32
รูปที่ 2.18 หลักการของระเบียบวิธีการเข้ารหัสแบบหน้าต่างเลื่อน.....	37
รูปที่ 2.19 พฤติกรรมของบิตในบัพเฟอร์ที่ตัวถอดรหัส.....	38
รูปที่ 2.20 การหาขนาดของบัพเฟอร์ตัวถอดรหัส และเวลาบรรจุก่อนจากบิตสตรีมวีดิทัศน์.....	40
รูปที่ 3.1 ภาพรวมของระบบ	43
รูปที่ 3.2 สัญญาณวีดิทัศน์ที่ใช้ในวิทยานิพนธ์.....	44
รูปที่ 3.3 จำนวนบิตเป้าหมาย และจำนวนบิตที่ลงรหัสจริงสำหรับลำดับภาพ Carphone.....	49

รูปที่ 3.4	ระเบียบวิธีการเข้ารหัสกับข้อกำหนดของขนาดของบัพเฟออร์ตัวถดถอย และเวลาบรรจุก่อน.....	51
รูปที่ 4.1	ค่าความแตกต่างสัมบูรณ์เฉลี่ยของลำดับวิถีที่คนที่ใช้ในการทดสอบ.....	55
รูปที่ 4.2	การวัดคุณภาพเชิงอัตวิสัยสำหรับลำดับวิถีที่คน Foreman ที่อัตราบิด เป้าหมาย 32 กิโลบิตต่อวินาที $W = 50$	57
รูปที่ 4.3	การวัดคุณภาพเชิงอัตวิสัยสำหรับลำดับวิถีที่คน Foreman ที่อัตราบิด เป้าหมาย 32 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W = 100$	58
รูปที่ 4.4	กราฟการเปรียบเทียบคุณภาพลำดับวิถีที่คน Foreman ที่อัตราบิดเป้าหมาย 32 กิโลบิตต่อวินาที	58
รูปที่ 4.5	การวัดคุณภาพเชิงอัตวิสัยสำหรับลำดับวิถีที่คน Foreman ที่อัตราบิด เป้าหมาย 64 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W = 50$	59
รูปที่ 4.6	การวัดคุณภาพเชิงอัตวิสัยสำหรับลำดับวิถีที่คน Foreman ที่อัตราบิด เป้าหมาย 64 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W = 100$	60
รูปที่ 4.7	กราฟการเปรียบเทียบคุณภาพลำดับวิถีที่คน Foreman ที่อัตราบิดเป้าหมาย 64 กิโลบิตต่อวินาที	61
รูปที่ 4.8	การวัดคุณภาพเชิงอัตวิสัยสำหรับลำดับวิถีที่คน Foreman ที่อัตราบิด เป้าหมาย 128 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W = 50$	62
รูปที่ 4.9	การวัดคุณภาพเชิงอัตวิสัยสำหรับลำดับวิถีที่คน Foreman ที่อัตราบิด เป้าหมาย 128 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W = 100$	63
รูปที่ 4.10	กราฟการเปรียบเทียบคุณภาพลำดับวิถีที่คน Foreman ที่อัตราบิดเป้าหมาย 128 กิโลบิตต่อวินาที	63
รูปที่ 4.11	การวัดคุณภาพเชิงอัตวิสัยสำหรับลำดับวิถีที่คน Foreman ที่อัตราบิด เป้าหมาย 256 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W = 50$	64
รูปที่ 4.12	การวัดคุณภาพเชิงอัตวิสัยสำหรับลำดับวิถีที่คน Foreman ที่อัตราบิด เป้าหมาย 256 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W = 100$	65
รูปที่ 4.13	กราฟการเปรียบเทียบคุณภาพลำดับวิถีที่คน Foreman ที่อัตราบิดเป้าหมาย 256 กิโลบิตต่อวินาที	65

รูปที่ 4.28	กราฟการเปรียบเทียบคุณภาพลำดับวิดิทัศน์ Suzie+Trevor ที่อัตราบิต เป้าหมาย 32 กิโลบิตต่อวินาที.....	82
รูปที่ 4.29	การวัดคุณภาพเชิงอัตตวิสัยสำหรับลำดับวิดิทัศน์ Suzie+Trevor ที่อัตราบิต เป้าหมาย 64 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W = 50$	83
รูปที่ 4.30	การวัดคุณภาพเชิงอัตตวิสัยสำหรับลำดับวิดิทัศน์ Suzie+Trevor ที่อัตราบิต เป้าหมาย 64 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W = 100$	84
รูปที่ 4.31	กราฟการเปรียบเทียบคุณภาพลำดับวิดิทัศน์ Suzie+Trevor ที่อัตราบิต เป้าหมาย 64 กิโลบิตต่อวินาที.....	84
รูปที่ 4.32	การวัดคุณภาพเชิงอัตตวิสัยสำหรับลำดับวิดิทัศน์ Suzie+Trevor ที่อัตราบิต เป้าหมาย 128 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W = 50$	85
รูปที่ 4.33	การวัดคุณภาพเชิงอัตตวิสัยสำหรับลำดับวิดิทัศน์ Suzie+Trevor ที่อัตราบิต เป้าหมาย 128 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W = 100$	86
รูปที่ 4.34	กราฟการเปรียบเทียบคุณภาพลำดับวิดิทัศน์ Suzie+Trevor ที่อัตราบิต เป้าหมาย 128 กิโลบิตต่อวินาที.....	86
รูปที่ 4.35	การวัดคุณภาพเชิงอัตตวิสัยสำหรับลำดับวิดิทัศน์ Suzie+Trevor ที่อัตราบิต เป้าหมาย 256 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W = 50$	88
รูปที่ 4.36	การวัดคุณภาพเชิงอัตตวิสัยสำหรับลำดับวิดิทัศน์ Suzie+Trevor ที่อัตราบิต เป้าหมาย 256 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W = 100$	89
รูปที่ 4.37	กราฟการเปรียบเทียบคุณภาพลำดับวิดิทัศน์ Suzie+Trevor ที่อัตราบิต เป้าหมาย 256 กิโลบิตต่อวินาที.....	89
รูปที่ 4.38	กราฟการเปรียบเทียบคุณภาพลำดับวิดิทัศน์ Suzie+Trevor ที่ขนาดหน้าต่าง $W=50$ และ 100 เมื่อมีข้อกำหนดของเวลาบรจุก่อน และขนาดของบัฟเฟอร์ ตัวถอดรหัส.....	93

บัญชีคำศัพท์

Analog	แอนะล็อก
Advanced Prediction Mode	โหมดประมาณการเคลื่อนไหวก้าวสูง
Bandwidth	แบนด์วิดท์
Bit Allocation	การจัดสรรบิต
Block – based Coding	การเข้ารหัสเชิงบล็อก
Block Layer	ชั้นบล็อก
Buffer	บัฟเฟอร์
Buffer Delay	การประวิงเวลาของบัฟเฟอร์
Buffer Fullness	ความเต็มของบัฟเฟอร์
Buffer Overflow	การล้นของบัฟเฟอร์
Buffer Underflow	การไม่มีข้อมูลในบัฟเฟอร์
Crominance Component	ส่วนประกอบความเข้มสี
Digital Video	วีดิทัศน์
Discrete Cosine Transform (DCT)	การทรานส์ฟอร์มแบบ DCT
Distortion	ความเพี้ยน
Dynamic Programming	การโปรแกรมพลวัต
Delay	การประวิงเวลา
Fixed Rate Channel	อัตราของสัญญาณแบบตายตัว
Frame Skipping	การกระโดดข้ามเฟรม
Global Motion Compensation	การชดเชยการเคลื่อนที่โดยรวม
Group of Blocks Layer	ชั้นกลุ่มบล็อก
Half – pixel Motion Estimation	การประมาณการเคลื่อนไหวกึ่งจุด
Information Technology (IT)	เทคโนโลยีสารสนเทศ
Intra frame coding	การเข้ารหัสในเฟรม
Inter frame coding	การเข้ารหัสระหว่างเฟรม

Lagrange	ลากรองจ์
Luminance Component	ส่วนประกอบความเข้มแสง
Macroblock Layer	ชั้นมาโครบล็อก
Mean	ค่าเฉลี่ย
Mean Absolute Different	ค่าความแตกต่างสัมบูรณ์เฉลี่ย
Memory	หน่วยความจำ
Model	แบบจำลอง
Motion Compensation	การชดเชยการเคลื่อนที่
Motion Estimation	การประมาณการเคลื่อนที่
Motion Vector	เวกเตอร์การเคลื่อนที่
Multimedia	มัลติมีเดีย
Object – based Coding	การเข้ารหัสเชิงวัตถุ
Offline	ออฟไลน์
Peak Signal to Noise Ratio (PSNR)	ค่าอัตราส่วนสัญญาณยดต่อสัญญาณรบกวน
Picture Layer	ชั้นรูปภาพ
Pixel	จุดภาพ
Pre-loading Time	เวลาบรรจุก่อน
Quantization Level	ระดับควอนไทซ์
Rate Control	การควบคุมอัตรา
Real - time Video Communication	การสื่อสารวิดีโอที่ค้นเวลาจริง
Sliding Window	หน้าต่างเลื่อน
Spatial Domain	โดเมนปริภูมิ โดเมนพื้นที่
Spatial Redundancy	ส่วนเกินในโดเมนปริภูมิ
Spatia l- temporal Correlation	ความสัมพันธ์ในเชิงพื้นที่และเวลา
Standard Deviation	ค่าเบี่ยงเบนมาตรฐาน
Streaming Video	วิดีโอที่ส่งสตรีมมิง
Test Sequence	ลำดับภาพมาตรฐาน
Unrestricted Motion Vector Mode	โหมดเวกเตอร์การเคลื่อนที่แบบไม่จำกัด

Variable Length Coding (VLC)	การเข้ารหัสแบบปรับเปลี่ยนความยาว
Video Coding	การเข้ารหัสสัญญาณวิดีโอ
Videoconference	การประชุมสัมมนาทางโทรทัศน์
Videophone	โทรศัพท์ภาพ
Wireless Channel	ช่องสัญญาณแบบไร้สาย
Zigzag Scan	การกวาดแบบซิกแซก



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 1

บทนำ

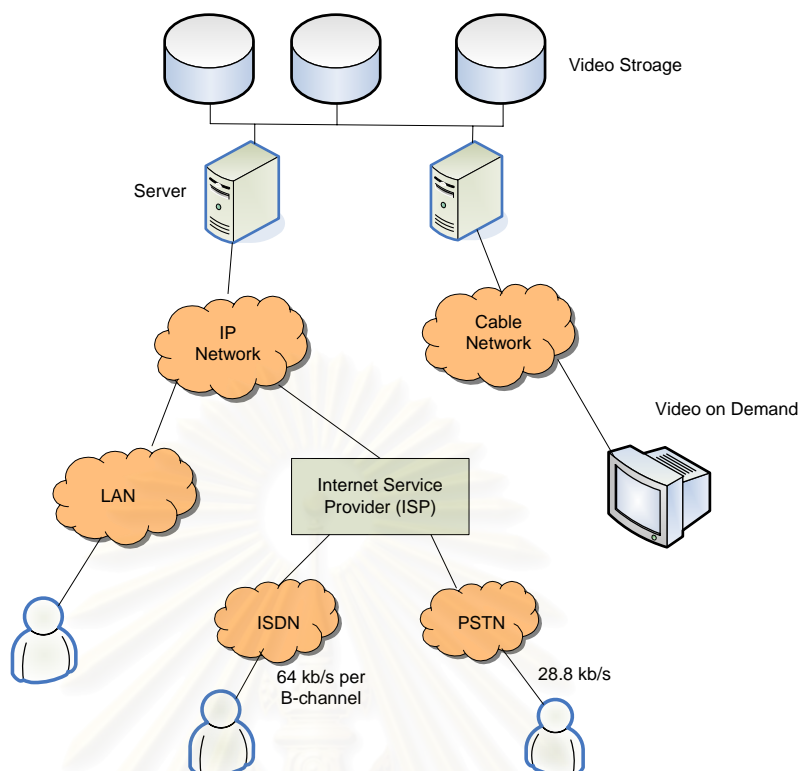
1.1 แนวเหตุผลและความเป็นมา

งานประยุกต์วีดิทัศน์ดิจิทัลเริ่มเข้ามามีบทบาทกับชีวิตของเรามากขึ้น ปัจจุบันมีมาตรฐานวีดิทัศน์ที่ถูกตั้งขึ้นมาเพื่อจุดประสงค์ที่แตกต่างกัน ตัวอย่างเช่น MPEG-1 [1] และ MPEG-2 [2] สำหรับงานประยุกต์สื่อประสม H.263 [3] [4] สำหรับงานประยุกต์การประชุมทางวีดิทัศน์ มาตรฐานทั้งหมดที่กล่าวมามีส่วนประกอบหนึ่งที่สำคัญในการเข้ารหัสสัญญาณวีดิทัศน์ คือการควบคุมอัตรา (Rate control) โดยการควบคุมอัตราจะทำการหาค่าควอนไทซ์พารามิเตอร์ที่เหมาะสม โดยคำนึงถึงจำนวนบิตสะสมในบัฟเฟอร์ แบนด์วิดท์ของช่องสัญญาณ และความซับซ้อนของเฟรม ซึ่งจะมีผลต่อค่าคุณภาพของสัญญาณวีดิทัศน์ที่ถูกเข้ารหัสโดยตรง ซึ่งในแต่ละมาตรฐานก็จะมีระเบียบวิธีการควบคุมอัตราที่แตกต่างกันไปขึ้นอยู่กับแบบจำลองทดสอบมาตรฐาน (test model) ที่ใช้ในมาตรฐานการเข้ารหัสสัญญาณวีดิทัศน์นั้น ๆ ยกตัวอย่างเช่น ในมาตรฐาน MPEG2 จะใช้ Test Model 5 (TM5) [5] และมาตรฐาน H.263 จะใช้ Test Model Near Term 8 (TMN8) [6]

ระเบียบวิธีการควบคุมอัตราจะมีหน้าที่จัดสรรบิตในการเข้ารหัสสัญญาณวีดิทัศน์ โดยการเลือกค่าควอนไทซ์พารามิเตอร์ที่เหมาะสมกับจำนวนบิตที่กำหนดมาให้เพื่อให้สัญญาณวีดิทัศน์ที่ถูกเข้ารหัสมีความผิดเพี้ยนน้อยที่สุด โดยเงื่อนไขบังคับอื่น ๆ เช่น สถานะของช่องสัญญาณ และค่าประวิงเวลาที่จำกัดจะถูกนำมาใช้ในการตัดสินใจในการเข้ารหัส [7] เพื่อให้ได้คุณภาพวีดิทัศน์ที่สูงที่สุด การหาอัตราความผิดเพี้ยนที่เหมาะสมที่สุด (rate-distortion optimization) จะถูกนำมาใช้หาค่าสัมประสิทธิ์ที่เหมาะสม ค่าควอนไทซ์ที่เหมาะสม [8] [9] เวกเตอร์การเคลื่อนที่ที่เหมาะสม [10] [11] และการเลือกโหมดในการเข้ารหัสที่เหมาะสม [12] ซึ่งเทคนิคที่นิยมใช้ในการจัดสรรบิตที่เหมาะสมที่ให้ค่าความผิดเพี้ยนน้อยที่สุดสำหรับการเข้ารหัสสัญญาณวีดิทัศน์ภายใต้เงื่อนไขบิตที่กำหนดให้ คือทฤษฎีลากรางจ์ (Lagrange theory) และการโปรแกรมพลวัต (Dynamic programming) [14] สิ่งหนึ่งที่ยากของเทคนิคนี้ คือต้องการการคำนวณที่สูงโดยเฉพาะกับการเข้ารหัสที่มีตัวเลือกของการเลือกโหมดในการเข้ารหัส ควอนไทซ์ และเวกเตอร์การเคลื่อนที่

ประเภทอื่น ๆ ของระเบียบวิธีการควบคุมอัตราจะถูกออกแบบโดยใช้แบบจำลองทางคณิตศาสตร์ [15]-[19] โดยใช้แบบจำลองอัตราและความผิดเพี้ยน (rate-distortion model) ในการหาระดับควอนไทซ์สำหรับการควบคุมอัตราบิดจากข้อมูล เช่นระดับการเติมของบัพเฟอร์ และค่าความแปรปรวนของค่าความผิดพลาดการทำนายที่ตกค้าง (variance of residual prediction-error) โดยในปัจจุบันการควบคุมอัตราด้วยวิธีนี้เป็นที่นิยมเนื่องจากไม่ซับซ้อน ข้อดีคือแบบจำลองทางคณิตศาสตร์ที่ถูกออกแบบมาไม่สามารถครอบคลุมกับลำดับวิธีทัศน์ได้ทุกประเภท เนื่องจากลำดับวิธีทัศน์มีมากมายหลากหลายประเภท ฉะนั้นถ้าแบบจำลองไม่เหมาะสมกับวิธีทัศน์เข้า คุณภาพของสัญญาณวิธีทัศน์จะไม่ดีนัก

ระเบียบวิธีการควบคุมอัตราบิดที่กล่าวมาแล้วถูกออกแบบมาสำหรับใช้กับสถานการณ์การสื่อสารในเวลาจริง ซึ่งการประวิงเวลาเป็นประเด็นสำคัญสำหรับการสื่อสารในเวลาจริง สำหรับในงานประยุกต์วิธีทัศน์สตรีมมิงสามารถแบ่งออกเป็นสองประเภท คือแบบออนไลน์ และแบบออฟไลน์ โดยในแบบออนไลน์ ลำดับวิธีทัศน์จะถูกเข้ารหัสเพียงรอบเดียว (one-pass encoding) เพื่อให้ทันส่งในเวลาจริง การประวิงเวลาจึงเป็นประเด็นสำคัญ แต่สำหรับงานประยุกต์วิธีทัศน์สตรีมมิงแบบออฟไลน์ เช่น วิธีทัศน์ตามคำขอ (Video-on-Demand) ห้องสมุดดิจิทัล (Digital Library) และการเรียนทางไกลแบบไม่ได้ตอบ (Non-Interactive Distance Learning) สำหรับงานประยุกต์เหล่านี้ลำดับวิธีทัศน์จะถูกเข้ารหัสล่วงหน้า และถูกเก็บไว้ในเครื่องคอมพิวเตอร์แม่ข่าย ผู้ใช้สามารถที่จะเข้าถึงข้อมูลได้ผ่านช่องสัญญาณที่มีอัตราบิดคงที่ เช่นโครงข่ายโทรศัพท์สาธารณะ (Public Switch Telephone Network) หรือโครงข่ายบริการสื่อสารร่วมระบบดิจิทัล (Integrated Services Digital Network) ดังแสดงในรูปที่ 1.1 เราจึงไม่ต้องพิจารณาเวลาประวิงของตัวเข้ารหัส ดังนั้นเราสามารถออกแบบระเบียบวิธีการควบคุมอัตราที่อาจจะต้องการการประมวลผลที่มากกว่าแต่สามารถให้คุณภาพของวิธีทัศน์ที่ดีกว่าได้ และประเด็นสำคัญที่เราต้องนำมาพิจารณาสำหรับงานประยุกต์วิธีทัศน์สตรีมมิงแบบออฟไลน์นี้ คือเวลาบรรจุก่อนในบัพเฟอร์ตัวถอดรหัส (pre-loaded time) ซึ่งหมายถึงเวลาที่เมื่อลำดับวิธีทัศน์มาถึงปลายทางจะถูกเก็บไว้ที่บัพเฟอร์ของตัวถอดรหัสก่อนที่จะถูกถอดรหัส เพื่อให้แน่ใจว่าทุกเฟรมจะถูกถอดรหัส และแสดงตรงตามเวลาที่ปลายทาง ถ้าไม่มีวิธีการที่เหมาะสมในการบรรจุก่อนการถอดรหัสจะไม่ถูกต้องตามเวลาเพราะว่าจะไม่มีบิตเข้ามาที่บัพเฟอร์ตัวถอดรหัสเนื่องมาจากช่องสัญญาณมีขนาดที่จำกัด



รูปที่ 1.1 ผู้ใช้สามารถเข้าถึงฐานข้อมูลวิดีโอที่สนับสนุนโครงข่ายสำหรับงานประยุกต์วิดีโอที่สนับสนุน

ในหลายปีที่ผ่านมา มีการพัฒนาเกี่ยวกับงานประยุกต์วิดีโอที่สนับสนุนแบบออนไลน์อย่างต่อเนื่อง ซึ่งจากการศึกษาและรวบรวม สามารถประมวลผลงานวิจัยที่เกี่ยวข้องได้ โดยเริ่มจาก Yue Yu และคณะ [32] ได้นำเสนอการเข้ารหัสวิดีโอที่สนับสนุนมาตรฐาน MPEG-2 โดยจะเข้ารหัสสองรอบ (two-pass encoding) โดยในรอบแรกของการเข้ารหัสจะทำการหาความสัมพันธ์ระหว่างค่าคอนโทปพารามิเตอร์ที่เป็นไปได้กับจำนวนบิตที่ได้จากการเข้ารหัสด้วยค่าคอนโทปพารามิเตอร์เหล่านั้น (R-Q function) ทำให้สามารถหาค่าคอนโทปพารามิเตอร์ที่เหมาะสมในแต่ละเฟรมได้ ซึ่งจะนำไปใช้ในการเข้ารหัสในรอบที่สอง ซึ่งวิธีนี้จะมีความซับซ้อนสูงเนื่องจากต้องทำการหาค่าคอนโทปพารามิเตอร์ที่เป็นไปได้ทั้งหมดในการเข้ารหัส I-Ming Pao และ Ming-Ting Sun [29] ได้นำเสนอระเบียบวิธีการควบคุมอัตราแบบหน้าต่างเลื่อนบนพื้นฐานของมาตรฐาน H.263 ซึ่งเป็นการใช้ข้อมูลทางสถิติของเฟรมในอนาคต คือค่าความแปรปรวนของเฟรม และจำนวนบิตของข้อมูลส่วนหัว มาช่วยในการจัดสรรบิต บิตจะถูกเก็บจากเฟรมที่ง่ายต่อการเข้ารหัส และนำไปใช้ในการเข้ารหัสกับเฟรมที่ต้องการบิตเพิ่มเพื่อให้คุณภาพของวิดีโอสูงขึ้น โดยเวลาบรรจุก่อน และขนาดบัพเฟอร์ตัวถอดรหัส ซึ่งเป็นพารามิเตอร์ที่สำคัญจะถูกนำมาพิจารณาด้วย Jianfei Cai และคณะ [33] ได้นำเสนอระเบียบวิธีการควบคุมอัตราสำหรับการเข้ารหัสวิดีโอที่สนับสนุนแบบออนไลน์ ซึ่งจะประยุกต์ระเบียบวิธีการควบคุมอัตราแบบหน้าต่างเลื่อนมาใช้ โดยการนำ

แบบจำลองอัตราและความผิดเพี้ยนโดเมน ρ (ρ -domain rate-distortion model) บนพื้นฐานของมาตรฐาน H.263 มาใช้ในการจัดสรรบิตให้เหมาะสมในระดับเฟรม

เนื่องจากมาตรฐานการเข้ารหัสสัญญาณวิดีโอ H.264 มีประสิทธิภาพในการเข้ารหัสที่สูงกว่าเมื่อเทียบกับมาตรฐานก่อนหน้า งานประยุกต์วิดีโอที่จะได้ประโยชน์จากมาตรฐาน H.264 ที่ออกมานี้จะรวมไปถึง วิดีทัศน์แบบไร้สาย (wireless video) วิดีทัศน์สตรีมมิง (streaming video) และอื่นๆ อีกมาก และประกอบกับงานวิจัยที่ผ่านมายังไม่มีการศึกษาเกี่ยวกับการควบคุมอัตราสำหรับการเข้ารหัสวิดีโอสตรีมมิงแบบออนไลน์บนมาตรฐาน H.264 มาก่อน ในวิทยานิพนธ์ฉบับนี้จึงได้นำเสนอเทคนิคการควบคุมอัตราสำหรับการเข้ารหัสวิดีโอสตรีมมิงแบบออนไลน์บนพื้นฐานของมาตรฐาน H.264 โดยอาศัยข้อมูลทางสถิติของเฟรมในขนาดตามาช่วยจัดสรรบิตให้กับวิดีโอ โดยจะทำการเข้ารหัสสองรอบ (two-pass encoding) ทำให้คุณภาพโดยรวมของวิดีโอดีขึ้น โดยจะพิจารณาถึงเงื่อนไขของเวลาบรจุก่อน และขนาดของบัฟเฟอร์ตัวถอดรหัส ซึ่งเป็นเงื่อนไขที่จำเป็นในงานประยุกต์วิดีโอสตรีมมิงแบบออนไลน์ด้วย

1.2 มาตรฐานการเข้ารหัสสัญญาณวิดีโอ

มาตรฐานการเข้ารหัสสัญญาณวิดีโอได้รับการพัฒนาอย่างต่อเนื่องในช่วง 20 ปีที่ผ่านมาโดยมาตรฐานที่รู้จักกันอย่างแพร่หลายได้รับการจัดทำโดย 2 องค์กรหลัก ได้แก่ ITU-T (International Telecommunication Union – Telecommunication Sector) และ ISO/IEC (International Standard Organization) โดยได้กำหนดมาตรฐานการเข้ารหัสสัญญาณวิดีโอดังนี้

1.2.1 มาตรฐานของ ITU-T

องค์กร ITU-T เริ่มต้นตั้งแต่ปลายทศวรรษที่ 1980 มุ่งเน้นการกำหนดมาตรฐานในการส่งข้อมูลวิดีโอที่ถูกบีบอัดผ่านโครงข่าย ใช้งานกับเครื่องโทรศัพท์ภาพ หรือการประชุมสัมมนาทางวิดีโอ โดยกำหนดมาตรฐานตระกูล H (H-series) มาตรฐานต่างๆ ในตระกูล H แสดงไว้ดังตารางที่ 1.1

ตารางที่ 1.1 มาตรฐานการเข้ารหัสสัญญาณวิดีโอทัศน์ของ ITU-T

หมายเลขมาตรฐาน	ปีที่กำหนดใช้	วัตถุประสงค์
H.261	ค.ศ.1990	เพื่อจัดการระบบการประชุมสัมมนาทางวิดีโอทัศน์ผ่านโครงข่าย ISDN (ISDN Videoconference network)
H.262 (ร่วมกับ ISO/IEC)	ค.ศ.1995	เพื่อการส่งกระจายสัญญาณ broadcast และพัฒนาการเก็บข้อมูลเหมือน MPEG-2
H.263	ค.ศ.1996	เพื่อจัดการและเข้ารหัสสัญญาณวิดีโอทัศน์เพื่อส่งผ่านโครงข่ายอัตราบิตต่ำและโครงข่ายไร้สาย
H.263+	ค.ศ.1998	พัฒนาจาก H.263 โดยมีฟังก์ชันการทำงานมากขึ้น เช่น การเข้ารหัสเพื่อป้องกันความผิดพลาด
H.263++	ค.ศ.2000	พัฒนาจาก H.263+ โดยมีฟังก์ชันการทำงานมากขึ้น เช่น ขนาดบล็อกสำหรับการประมาณการเคลื่อนที่มีขนาดเล็กลง
H.26L	ค.ศ.2002	เพื่อจัดการและเข้ารหัสสัญญาณวิดีโอทัศน์เพื่อส่งผ่านโครงข่ายอินเทอร์เน็ตและการให้บริการสัญญาณวิดีโอทัศน์

1.2.2 มาตรฐานของ ISO/IEC

มาตรฐาน MPEG เป็นมาตรฐานซึ่งตั้งขึ้นโดยคณะกรรมการ MPEG (Moving Picture Experts Group) โดยเริ่มขึ้นในปี ค.ศ. 1988 โดยมีวัตถุประสงค์ปีบอัดสัญญาณวิดีโอทัศน์และออกดีโอ เพื่อเก็บลงสื่อดิจิทัล และส่งผ่านโครงข่ายอัตราบิตสูง โดยมาตรฐานต่าง ๆ แสดงดังตารางที่ 1.2

ตารางที่ 1.2 มาตรฐานการเข้ารหัสสัญญาณวิดีโอทัศน์ของ ISO/IEC

หมายเลขมาตรฐาน	ปีที่กำหนดใช้	วัตถุประสงค์
MPEG-1	ค.ศ.1992	เพื่อเข้ารหัสข้อมูลวิดีโอทัศน์สำหรับสื่อประสมและการเก็บข้อมูลในสื่อเก็บข้อมูล เช่น แผ่นซีดีรอม (Video CD) และ MP3

MPEG-2	ค.ศ.1995	พัฒนาจาก MPEG-1 เพื่อการส่งกระจายสำหรับโทรทัศน์ดิจิทัลความละเอียดสูง(HDTV) โทรทัศน์ผ่านดาวเทียมและเคเบิล รวมถึงพัฒนาการเก็บข้อมูลลงสื่อเก็บข้อมูลที่มีความจุมากขึ้น เช่น DVD (Digital Video Disc)
MPEG-4	ค.ศ. 1999 (เวอร์ชัน 1) ค.ศ. 2000 (เวอร์ชัน 2)	เพื่อการส่งสัญญาณวีดิทัศน์แบบอัตราบิตต่ำมาก (Very low bit rate) ถึงสูงมาก และนำไปประยุกต์ใช้งานทางอินเทอร์เน็ตแบบโต้ตอบกันได้ (Interactive) มุ่งเน้นในการบีบอัดและใช้งานทางสื่อประสม
MPEG-7	ค.ศ. 2001	เพื่อพัฒนามาตรฐานการสืบค้นหาและการจัดเก็บข้อมูลในรูปแบบของสื่อประสม
MPEG-21	ค.ศ.2002	Multimedia Framework เพื่อวางระบบการสื่อสารสื่อประสมให้สามารถสื่อสารกันเป็นระบบเดียวกัน

สำหรับมาตรฐานการเข้ารหัสสัญญาณวีดิทัศน์ H.264 เริ่มต้นเมื่อในปี ค.ศ. 1998 ได้ริเริ่มพัฒนามาตรฐานใหม่โดยเบื้องต้นใช้ชื่อว่า H.26L โดยต่อมาเปลี่ยนชื่อเป็น H.264 หรือ AVC โดยร่วมมือกันระหว่าง ITU-T และ ISO/IEC เป็น Joint video team (JVT) โดย ITU-T ออกเป็นมาตรฐานภายใต้ชื่อว่า H.264 และ ISO/IEC ออกเป็นมาตรฐานภายใต้ชื่อว่า MPEG-4 PART 10 (Advanced Video Coding: AVC) วัตถุประสงค์เพื่อให้เหมาะสมสำหรับการประยุกต์ใช้กับงานทางวีดิทัศน์ดิจิทัลแบบเต็มรูปแบบ [20] เช่น การสื่อสารสัญญาณวีดิทัศน์ที่มีอัตราบิตต่ำสำหรับโครงข่ายไร้สาย การแพร่ภาพบนโทรทัศน์มาตรฐาน (standard-definition television) การแพร่ภาพบนโทรทัศน์ความคมชัดสูง (high definition broadcast television) การสตรีม (stream) สัญญาณวีดิทัศน์ผ่านอินเทอร์เน็ต (internet) และการแสดงภาพคุณภาพสูงสุดสำหรับการประยุกต์ใช้ในโรงภาพยนตร์ เป็นต้น

1.3 แนวทางที่นำเสนอ

เนื่องจากในปัจจุบันงานวิจัยทางการเข้ารหัสสัญญาณวีดิทัศน์ มีแนวโน้มของการพัฒนาการวิจัยทางการเข้ารหัสสัญญาณวีดิทัศน์เพื่อใช้งานในด้านการสื่อสารผ่านช่องสัญญาณอัตราบิตต่ำใช้สำหรับการสื่อสารเวลาจริง ซึ่งการประวิงเวลาเป็นปัจจัยสำคัญ สำหรับการสื่อสารที่ไม่ใช่เวลาจริง เช่น งานประยุกต์วีดิทัศน์สตรีมมิงแบบออฟไลน์ เงื่อนไขจะต่างออกไป เมื่อลำดับสัญญาณวีดิทัศน์ทำการเข้ารหัสออฟไลน์ จึงไม่ต้องพิจารณาเวลาประวิงของตัวเข้ารหัส เวลาบรรจุก่อนในบัพเฟอร์ตัวถอดรหัสก่อนที่จะมีการแสดงที่ปลายทาง และขนาดของบัพเฟอร์ตัวถอดรหัส เป็นสิ่งที่ต้องพิจารณา งานวิจัยที่ผ่านมาเป็นการพัฒนาลงบนมาตรฐานการเข้ารหัสสัญญาณวีดิทัศน์ H.263 และ MPEG-2 [29][32][33] เนื่องจากในมาตรฐาน H.264 มีประสิทธิภาพในการเข้ารหัสที่สูงกว่ามาตรฐานก่อนหน้า และในงานวิจัยที่ผ่านมา ยังไม่มีการศึกษาในมาตรฐานการเข้ารหัสสัญญาณวีดิทัศน์ H.264

ดังนั้นในวิทยานิพนธ์นี้จึงได้มีการศึกษาและพัฒนาเทคนิคการควบคุมอัตราสำหรับการเข้ารหัสวีดิทัศน์สตรีมมิงแบบออฟไลน์บนพื้นฐานของมาตรฐาน H.264 เพื่อให้คุณภาพของสัญญาณวีดิทัศน์ดีขึ้น เนื่องจากการเข้ารหัสแบบออฟไลน์เราสามารถจะประยุกต์ใช้ระเบียบวิธีการเข้ารหัสด้วยหน้าต่างเลื่อน (Sliding-Window) [29][33] ซึ่งเป็นกรนำข้อมูลทางสถิติของเฟรมในอนาคตมาช่วยในการจัดสรรบิตทำให้คุณภาพของวีดิทัศน์ดีขึ้นได้ และวิธีของการประมาณบิตเป้าหมายแบบปรับตัวได้ตามความซับซ้อนของการเข้ารหัสในแต่ละเฟรม โดยการใช้อัตราส่วนค่าความแตกต่างสัมบูรณ์เฉลี่ย (MAD ratio) ในการวัดความซับซ้อนของเฟรม [30] ซึ่งเทคนิคนี้จะนำมาใช้ร่วมกับระเบียบวิธีการเข้ารหัสด้วยหน้าต่างเลื่อน โดยเทคนิคนี้สามารถลดการตกลงอย่างฉับพลันของคุณภาพของวีดิทัศน์ในช่วงที่มีการเปลี่ยนของฉาก และเฟรมที่มีการเคลื่อนที่มาก ๆ ได้

1.4 วัตถุประสงค์

พัฒนาระเบียบวิธีการควบคุมอัตราสำหรับการเข้ารหัสวีดิทัศน์สตรีมมิงแบบออฟไลน์ สำหรับมาตรฐานการเข้ารหัสสัญญาณวีดิทัศน์ H.264

1.5 ขอบเขตวิทยานิพนธ์

พัฒนาระเบียบวิธีการควบคุมอัตราบิดสำหรับงานประยุกต์วีดิทัศน์สตรีมมิงโดยใช้ระเบียบวิธีการเข้ารหัสหน้าต่างเลื่อนมาใช้ภายใต้เงื่อนไขของบัพเฟอรัต์วอลอตรหัสและเวลาบรจุก่อนสำหรับมาตรฐานการเข้ารหัสสัญญาณวีดิทัศน์ H.264

1.6 ขั้นตอนการดำเนินงาน

1. ศึกษาทฤษฎีการเข้ารหัสสัญญาณวีดิทัศน์มาตรฐาน H.263 และ H.264
2. ศึกษาระเบียบวิธีการควบคุมอัตราบิดที่ใช้ในแบบจำลองมาตรฐาน H.263 และ H.264
3. พัฒนาระเบียบวิธีควบคุมอัตราสำหรับการเข้ารหัสวีดิทัศน์สตรีมมิงแบบออฟไลน์สำหรับมาตรฐานการเข้ารหัสสัญญาณวีดิทัศน์ H.264
4. วิเคราะห์และเปรียบเทียบผลของวิธีที่นำเสนอ
5. สรุปและจัดทำรูปเล่มวิทยานิพนธ์

1.7 ประโยชน์ที่คาดว่าจะได้รับ

1. ได้ศึกษาความรู้พื้นฐานในการเข้ารหัสสัญญาณวีดิทัศน์และการส่งสัญญาณวีดิทัศน์
2. พัฒนาซอฟต์แวร์ และระเบียบวิธีการควบคุมอัตราในส่วนของ การควบคุมอัตราสำหรับมาตรฐานการเข้ารหัสสัญญาณวีดิทัศน์ H.264 เพื่อใช้เป็นข้อมูลในการควบคุมอัตราสำหรับการเข้ารหัสวีดิทัศน์สตรีมมิงแบบออฟไลน์บนพื้นฐานของมาตรฐาน H.264
3. สามารถนำไปประยุกต์ใช้งานกับการส่งสัญญาณวีดิทัศน์สตรีมมิงได้

บทที่ 2

ความรู้พื้นฐาน

เนื้อหาในบทนี้กล่าวถึงทฤษฎีต่างๆ ที่เกี่ยวข้องกับวิทยานิพนธ์ คือพื้นฐานของวีดิทัศน์ ดิจิทัล หลักการทำงานพื้นฐานของการเข้ารหัสสัญญาณวีดิทัศน์ การควบคุมอัตราสำหรับการเข้ารหัสสัญญาณวีดิทัศน์ H.264 การทบทวนงานวิจัยที่เกี่ยวข้อง

2.1 วีดิทัศน์ดิจิทัล (Digital Video)

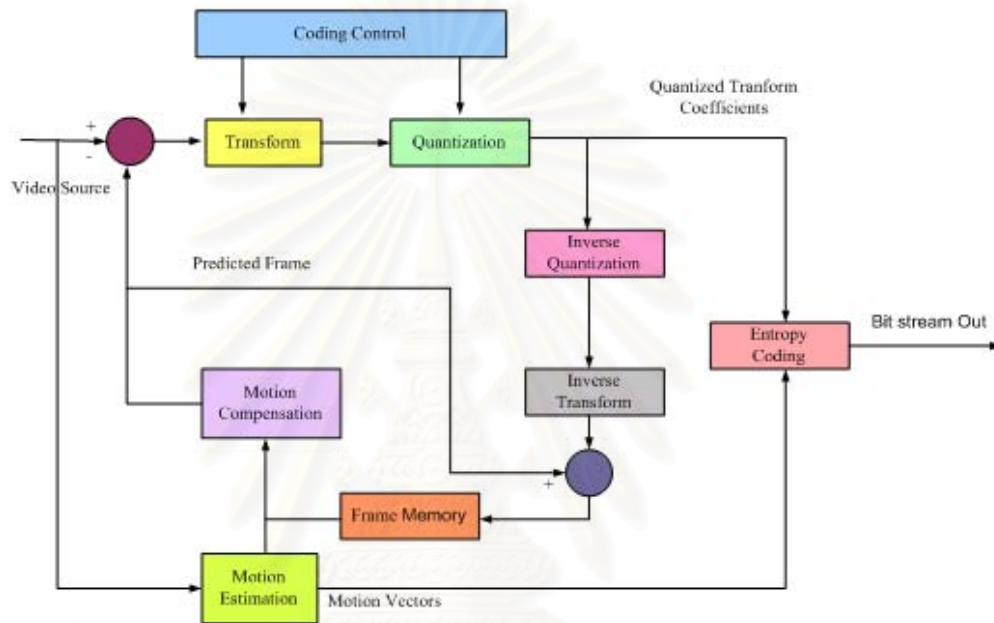
วีดิทัศน์ดิจิทัลเป็นกลุ่มภาพที่มาจากการบินที่จากกล้องดิจิทัล หรือสแกนรูปจากภาพแอนะล็อก (analog) โดยมีหน่วยเป็นเฟรมต่อหนึ่งหน่วยเวลา เช่น เฟรมต่อวินาที แต่ละภาพจะประกอบไปด้วยจุดภาพมาจากการสุ่มตัวอย่าง และผ่านการควอนไทซ์ โดยแต่ละจุดภาพประกอบด้วย ส่วนประกอบความส่องสว่าง (luminance component) และส่วนประกอบสี (chrominance component) โดยทั่วไปความส่องสว่างของแต่ละภาพเมื่อผ่านการควอนไทซ์แล้ว จะถูกแบ่งออกเป็น 256 ระดับ ดังนั้นจุดภาพแต่ละจุดสามารถแสดงด้วยความส่องสว่างตั้งแต่ 0 ถึง 255 สำหรับความเข้มสีไม่มีความจำเป็นต้องแสดงทุกจุดภาพเนื่องจากความสามารถในการรับสีของระบบประสาทในการมองเห็นของมนุษย์น้อยกว่าความส่องสว่างจึงมีจำนวนจุดความเข้มสีแล้วแต่กำหนด [21] รูปแบบของสัญญาณวีดิทัศน์บางชนิดกำหนดไว้ดังตารางที่ 2.1

ตารางที่ 2.1 รูปแบบของสัญญาณวีดิทัศน์ที่ใช้งานในปัจจุบัน

รูปแบบ	ความละเอียด	อัตราเฟรม	Y:Cb:Cr	การกวาด
CCIR601(NTSC)	720 x 480	30	4:2:2	interlaced
CCIR601(PAL)	720 x 576	25	4:2:2	interlaced
SIF(NTSC)	352 x 240	30	4:2:0	progressive
SIF(PAL)	352 x 288	25	4:2:0	progressive
CIF	352 x 288	30	4:2:0	progressive
QCIF	176 x 144	30	4:2:0	progressive

2.2 การเข้ารหัสสัญญาณวิดีโอ (Video Coding)

ในปัจจุบันมีการใช้มาตรฐานการเข้ารหัสสัญญาณวิดีโอที่ต่างกันออกไปอย่างแพร่หลาย ซึ่งแต่ละมาตรฐานมีวัตถุประสงค์ และการประยุกต์ใช้งานแตกต่างกันไป สำหรับการเข้ารหัสสัญญาณวิดีโอที่มีหลักการการทำงานหลักๆ แสดงดังรูปที่ 2.1



รูปที่ 2.1 บล็อกไดอะแกรมของการเข้ารหัสสัญญาณวิดีโอที่มีการบีบอัด

สัญญาณวิดีโอที่เข้าจะถูกรวบรวมให้อยู่ในระดับบล็อกที่มีขนาดเท่ากับ 16x16 จุดภาพ หรือเรียกว่ามาโครบล็อก แต่ละมาโครบล็อกสามารถเข้ารหัสภายในเฟรม (Intra frame coding) หรือ การเข้ารหัสระหว่างเฟรม (Inter frame coding) โดยเฟรมแรกจะถูกแปลงจากโดเมนทางเวลามาเป็นโดเมนความถี่ มาตรฐาน H.263 ใช้การแปลงแบบ DCT (Discrete Cosine Transform) ขนาดบล็อก 8x8 แต่สำหรับมาตรฐาน H.264 ใช้การแปลงแบบจำนวนเต็ม (Integer Transform) ขนาดบล็อก 4x4 การแปลงแบบจำนวนเต็มใช้พื้นฐานจากการแปลงแบบ DCT ซึ่งกรณีที่ทำอินเวอร์สการแปลงจะมีประสิทธิภาพที่ดีกว่ารูปแบบการแปลงแบบ DCT จากนั้นจะทำการควอนไทซ์ค่าสัมประสิทธิ์เพื่อบีบอัดข้อมูล จากนั้นสัมประสิทธิ์ที่ได้มาจะถูกนำจัดเรียงใหม่ให้อยู่ในรูปแบบมิติเดียวโดยการกราดแบบซิกแซก (zigzag scan) และทำการเข้ารหัสเอนโทรปี ในการเข้ารหัสสัญญาณวิดีโอ เมื่อผ่านการแปลง และการควอนไทซ์แล้วจะทำการแปลงย้อนกลับเพื่อให้อยู่ในรูปแบบจุดภาพเช่นเดิมเพื่อนำไปประมาณการเคลื่อนที่ โดยตัวเข้ารหัสจะค้นหาบล็อก

หน้าที่เก็บไว้ในหน่วยความจำว่าบล็อกใดที่มีค่าความแตกต่างระหว่างมาโครบล็อกของเฟรมปัจจุบันและเฟรมก่อนหน้าที่มีค่าน้อยที่สุด ทำให้ได้เวกเตอร์การเคลื่อนที่ (motion vector) เมื่อขดเซกการเคลื่อนที่ตัวเข้ารหัสจะเก็บสัญญาณวีดิทัศน์เฟรมใหม่เพื่อสามารถใช้สำหรับเป็นเฟรมอ้างอิงต่อไป

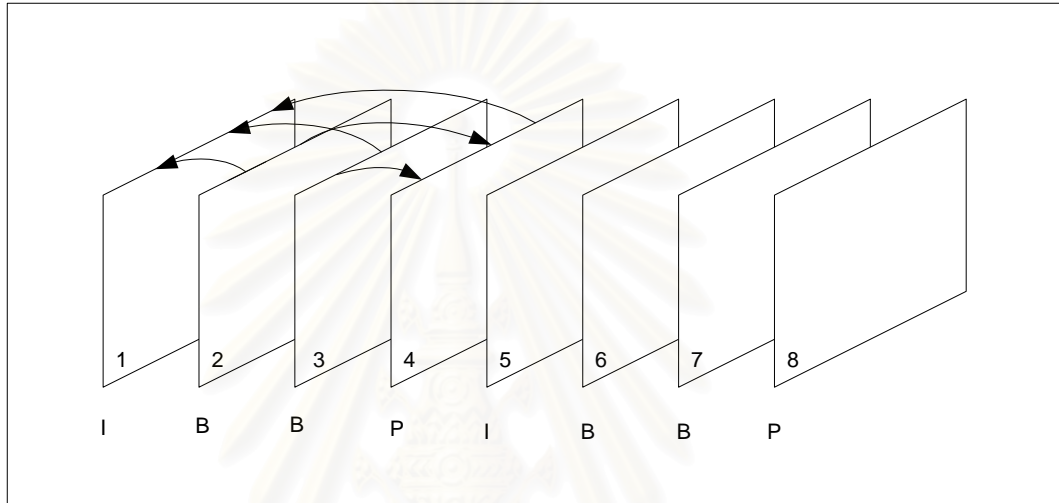
ในการเข้ารหัสสัญญาณวีดิทัศน์สามารถแบ่งได้ 2 ประเภท คือการเข้ารหัสภายในเฟรมเฟรมที่ถูกเข้ารหัสเรียกว่า I เฟรม (intra coding frame) และการเข้ารหัสระหว่างเฟรม (inter coding frame) เฟรมที่ถูกเข้ารหัสจะเรียกว่า P เฟรม (predictive coding frame) หรือ B เฟรม (Bidirectionally interpolated frame) ซึ่งจะกล่าวดังต่อไปนี้

2.2.1 ชนิดของรูปภาพ

ชนิดของรูปภาพแบ่งตามลักษณะของการเข้ารหัสสามารถแบ่งได้ 3 ชนิด ได้แก่ I และ B โดยมีรายละเอียดดังนี้

1. I (Intra) เป็นภาพที่เข้ารหัสโดยไม่ต้องอาศัยข้อมูลจากเฟรมอื่น ดังนั้นลำดับภาพของวีดิทัศน์เฟรมแรกจะเป็นภาพแบบ I เสมอ โดยเป็นเฟรมที่ถูกบีบอัดภายในเฟรม ใช้เป็นจุดเข้าถึง (access point) ของการเข้าถึงแบบสุ่ม ภาพแบบ I มีอัตราการบีบอัดข้อมูลต่ำที่สุดเมื่อเทียบกับภาพทั้งสามแบบ
2. P (Predicted) เป็นภาพที่เข้ารหัสโดยอาศัยข้อมูลของเฟรมก่อนหน้า และสามารถถอดรหัสโดยอาศัยข้อมูลของเฟรมก่อนหน้า ซึ่งเฟรมก่อนหน้าเรียกว่า เฟรมอ้างอิงซึ่งอาจจะเป็น I หรือ P เฟรมก็ได้ สำหรับการอ้างอิงถึงข้อมูลของเฟรมอ้างอิงต้องระบุเวกเตอร์การเคลื่อนที่มาด้วย ผลของการบีบอัดภาพแบบ P มีขนาดประมาณร้อยละ 30 ถึง 50 ของเฟรมแบบ I [34]
3. B (Bidirectionally interpolated) เป็นภาพที่เข้ารหัสและถอดรหัสโดยอาศัยข้อมูลทั้งจากอดีตและอนาคต รวมกับค่าการเคลื่อนที่ของเฟรมดังกล่าวทั้งสองเฟรม โดยละเลยค่าความเพี้ยนของภาพที่ได้จากการประมาณ ดังนั้นเฟรมชนิดนี้จะไม่ใช้เป็นเฟรมอ้างอิง ภาพชนิดนี้ให้การบีบอัดข้อมูลสูงสุด กล่าวคือ มีขนาดเหลือเพียงร้อยละ 15 ถึง 25 ของเฟรมแบบ I เท่านั้น [34]

โดยถ้าเรียงจำนวนบิตที่ต้องใช้สำหรับเข้ารหัสชนิดต่างจากมากไปหาน้อย พบว่ามากที่สุดคือภาพแบบ I รองลงมาคือภาพแบบ P น้อยที่สุด คือภาพแบบ B ความสัมพันธ์ระหว่างภาพแบบ I P และ B เป็นดังรูปที่ 2.2 ภาพเฟรมลำดับที่ 4 ซึ่งเป็นภาพแบบ P ได้จากการอ้างอิงเฟรมลำดับที่ 1 (เฟรมแบบ I) ส่วนเฟรมแบบ B ทั้งสองเฟรมในลำดับที่ 2 และ 3 ได้จากการอ้างอิงเฟรมแบบ I และ P



รูปที่ 2.2 ความสัมพันธ์ระหว่างเฟรมแบบ I P และ B ในการทำนายการเคลื่อนที่

2.2.2 การเข้ารหัสภายในเฟรม

การเข้ารหัสชนิดนี้เป็นการลดความซ้ำซ้อนในโดเมนปริภูมิ (Spatial Redundancy) ของเฟรมภาพด้วยการบีบอัดข้อมูลพื้นฐานการพิจารณาความซ้ำซ้อนของข้อมูลภายในหนึ่งเฟรมโดยเทคนิคต่างๆดังนี้

1. การควอนไทซ์แบบหยาบ (Coarse Quantization) เป็นการลดจำนวนข้อมูลภายในเฟรมโดยข้อมูลบางส่วนของภาพถูกทิ้งไป
2. การเข้ารหัสแบบการแปลง (Transform coding) ซึ่งเป็นการแปลงข้อมูลภาพจากโดเมนปริภูมิ (Spatial domain) เป็นโดเมนความถี่ (Frequency Domain) สัมประสิทธิ์จากการแปลงจะอยู่ในรูปสองมิติและถูกจัดให้อยู่มิติเดียวเพื่อทำการเข้ารหัสต่อไป

3. การทำควอนไทซ์แบบเวกเตอร์ (Vector Quantization) ข้อมูลภาพจะถูกแบ่งกลุ่ม และแสดงแทนด้วยสัญลักษณ์จากชุดรหัส (code book) เพื่อทำการส่งสัญลักษณ์แทนเท่านั้น ซึ่งเป็นการประหยัดจำนวนบิตในการส่ง

สำหรับมาตรฐานการเข้ารหัสสัญญาณวีดิทัศน์ H.264 [22] การเข้ารหัสภายในเฟรมแบ่งออกเป็นบล็อก ส่วนประกอบความส่องสว่างขนาด 4×4 จุดภาพทั้งหมด 9 โหมด และบล็อกส่วนประกอบความส่องสว่างขนาด 16×16 จุดภาพ ทั้งหมด 4 โหมด และสำหรับบล็อกส่วนประกอบสีขนาด 4×4 จุดภาพ 4 โหมด โดยการเลือกโหมดของการเข้ารหัสเลือกจากการหาค่าความแตกต่างสัมบูรณ์รวม (Sum Absolute Difference: SAD) ที่น้อยที่สุด ดังสมการที่ (2.1)

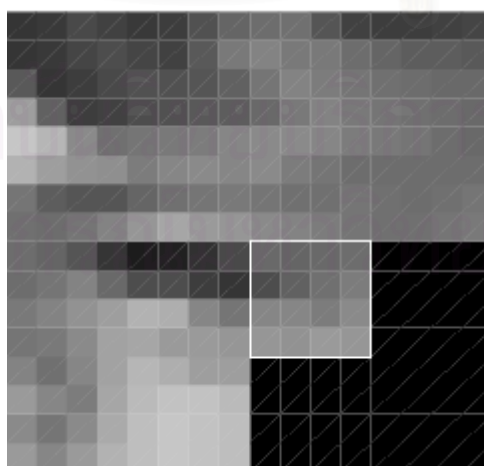
$$SAD = \sum_{i=1}^N \sum_{j=1}^N |I(i, j) - O(i, j)| \quad (2.1)$$

โดยที่ $N \times N$ เป็นขนาดบล็อก

$I(i, j)$ คือ ค่าส่วนประกอบความส่องสว่างที่ได้จากการประมาณ

$O(i, j)$ คือ ค่าส่วนประกอบความส่องสว่างดั้งเดิม

สำหรับบล็อกส่วนประกอบความส่องสว่างขนาด 4×4 จุดภาพที่จะทำการเข้ารหัสแสดงดังรูปที่ 2.3 โดยแต่ละจุดภาพ (a ถึง p) สามารถแสดงดังรูปที่ 2.3 โดยแต่ละโหนดมีรายละเอียดดังนี้



รูปที่ 2.3 ส่วนประกอบความส่องสว่างขนาด 4×4 จุดภาพ [20(22)]

M	A	B	C	D	E	F	G
I	a	b	c	d			
J	e	f	g	h			
K	i	j	k	l			
A	m	n	o	p			

รูปที่ 2.4 ตำแหน่งส่วนประกอบความส่องสว่างขนาด 4 x 4 จุดภาพ

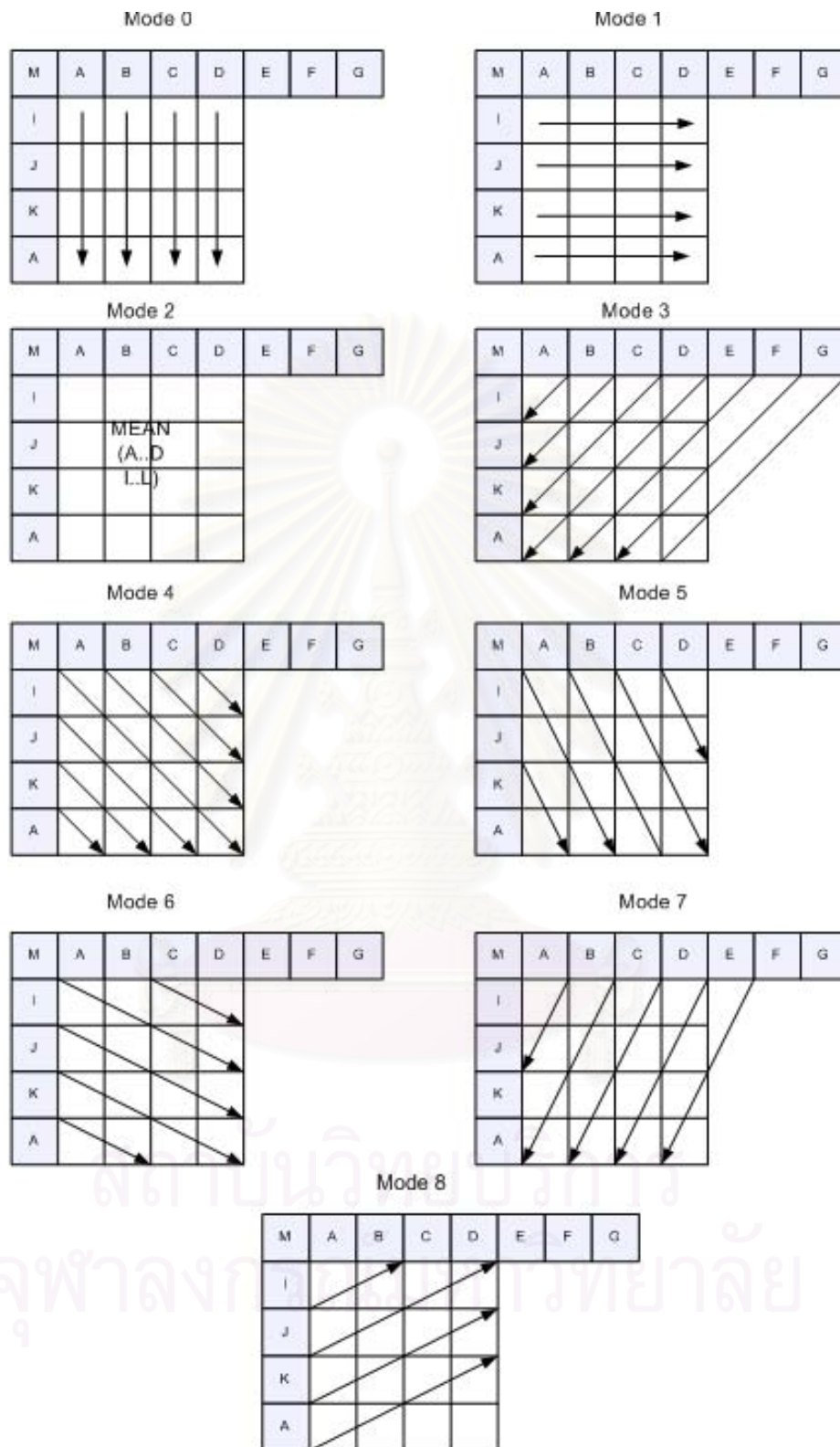
- โหมด 0 สามารถคำนวณจากการประมาณในแนวตั้งจากจุดภาพ A B C และ D
- โหมด 1 สามารถคำนวณจากการประมาณในแนวนอนจากจุดภาพ I J K และ L
- โหมด 2 สามารถคำนวณจากค่าเฉลี่ยตั้งแต่ A B C D และ I J K และ L
- โหมด 3 สามารถคำนวณจากแนวทแยงซ้ายที่ 45 องศา จากซ้ายลงมาทางขวา
- โหมด 4 สามารถคำนวณจากแนวทแยงขวาที่ 45 องศา จากขวาลงมาทางซ้าย
- โหมด 5 สามารถคำนวณที่ 26.6 องศาแนวตั้งทางขวาจากซ้ายลงมาทางขวา
- โหมด 6 สามารถคำนวณที่ 26.6 องศาต่ำกว่าแนวนอนจากซ้ายลงมาทางขวา
- โหมด 7 สามารถคำนวณที่ 26.6 องศาแนวตั้งทางขวาจากขวาลงมาทางซ้าย
- โหมด 8 สามารถคำนวณที่ 26.6 องศาสูงกว่าแนวนอนจากขวาลงมาทางซ้าย

โดยสามารถแสดงดังรูปที่ 2.5

สำหรับบล็อกส่วนประกอบความส่องสว่างขนาด 16x16 จุดภาพ สามารถหาได้จาก 4 โหมด รายละเอียดดังนี้

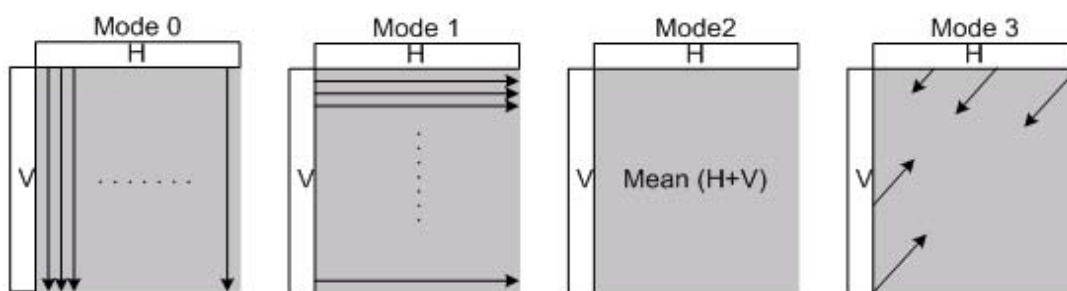
- โหมด 0 สามารถคำนวณจากการประมาณในแนวตั้งจากกลุ่มจุดภาพ H
- โหมด 1 สามารถคำนวณจากการประมาณในแนวตั้งจากกลุ่มจุดภาพ V
- โหมด 2 สามารถคำนวณจากค่าเฉลี่ยจากกลุ่มจุดภาพ H และ V
- โหมด 3 สามารถคำนวณจากกลุ่มจุดภาพ H ในแนวทแยงจากขวามาซ้าย และกลุ่มจากจุดภาพ V ในแนวทแยงจากซ้ายไปขวา

โดยสามารถแสดงดังรูปที่ 2.6



รูปที่ 2.5 การเข้ารหัสภายในเฟรมของส่วนประกอบความส่องสว่างขนาด 4x4 จุดภาพ

โหมด 0 ถึง 8



รูปที่ 2.6 การเข้ารหัสภายในเฟรมของส่วนประกอบความส่องสว่างขนาด 16x16 จุดภาพ
โหมด 0 ถึง 3

สำหรับการเข้ารหัสภายในเฟรมของส่วนประกอบความเข้มสีขนาด 8x8 จุดภาพ มีทั้งหมด 4 โหมดลักษณะคล้ายกับการเข้ารหัสภายในเฟรมของส่วนประกอบความส่องสว่างขนาด 16x16 จุดภาพ

2.2.3 การเข้ารหัสระหว่างเฟรม

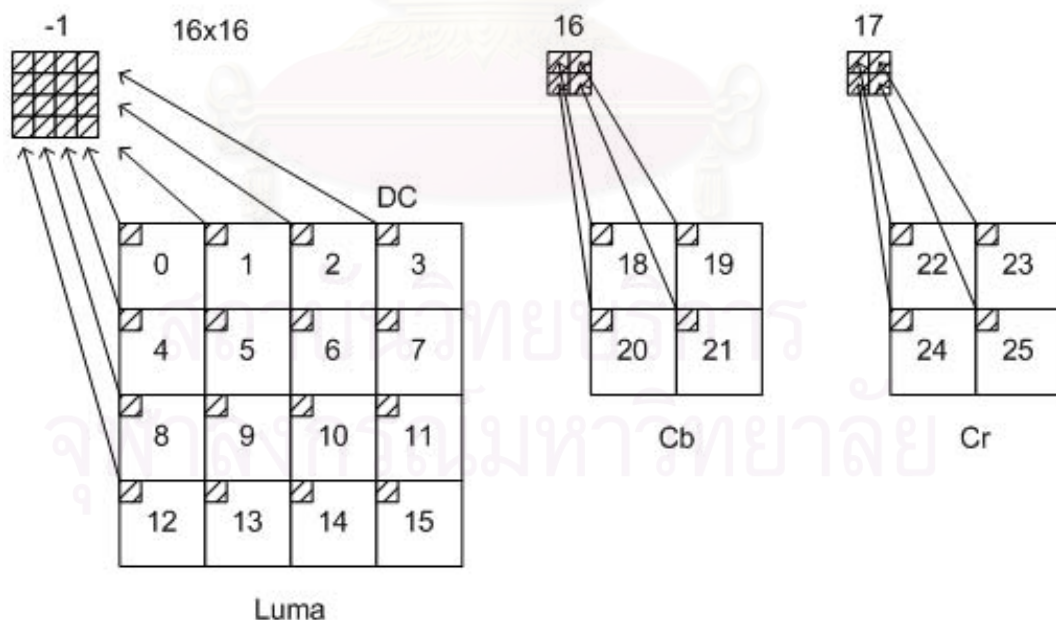
เนื่องจากสัญญาณวิดีโอที่ติดกันจะมีความสัมพันธ์ระหว่างข้อมูลทางปริภูมิ-เวลา (Spatial-Temporal correlation) ซึ่งทำให้เกิดความซ้ำซ้อนของข้อมูล เพื่อต้องการลดข้อมูลที่ซ้ำซ้อนนี้จึงใช้เทคนิคต่างๆ ดังนี้

1. การเข้ารหัสแบบสุ่มทางเวลา (Time Subsampling Coding) การเข้ารหัสเฉพาะบางเฟรม วิธีนี้ทำให้ลดข้อมูลได้มากแต่อาจทำให้ลำดับภาพขาดความต่อเนื่อง
2. การเข้ารหัสค่าความแตกต่าง (Different Coding) มี 2 แบบ คือการเข้ารหัสเฉพาะความแตกต่างระหว่างเฟรม และการเข้ารหัสค่าความแตกต่างระหว่างบล็อก (Block-Based Different Coding) โดยการเข้ารหัสเฉพาะความแตกต่างระหว่างเฟรมอาศัยความสัมพันธ์ของลำดับภาพที่อยู่ติดกันจึงมีความคล้ายคลึงกัน แต่อาจต้องใช้ส่วนหัว (overhead) ในการบ่งบอกว่าจุดภาพใดมีการเปลี่ยนแปลง จุดภาพใดไม่มีการเปลี่ยนแปลง ส่วนการเข้ารหัสค่าความแตกต่างระหว่างบล็อกต่างกันตรงที่ภาพจะถูกแบ่งเป็นบล็อกย่อยๆ ก่อน หลังจากนั้นแต่ละบล็อกจะถูกนำมาเปรียบเทียบ
3. การประมาณการเคลื่อนที่ (Motion Estimation) และการชดเชยการเคลื่อนที่ (Motion Compensation) เป็นวิธีการลดข้อมูลในการเข้ารหัส โดยเข้ารหัสเฉพาะความแตกต่างแทนที่จะเข้ารหัสทั้งเฟรม โดยทั่วไปเราสามารถสมมติว่าค่าของจุดภาพใน

ภาพที่มีการเคลื่อนที่ของวัตถุได้เลื่อนจากจุดหนึ่งในเฟรมหนึ่งไปยังจุดอื่นในเฟรมถัดไป ถ้าเราสามารถหาตำแหน่งที่เปลี่ยนแปลง หรือที่เรียกว่าเวกเตอร์การเคลื่อนที่ ทำให้สามารถลดความแตกต่างลงไปได้ วิธีนี้เรียกว่าการชดเชยการเคลื่อนที่ การประมาณการเคลื่อนที่จึงมีบทบาทสำคัญในส่วนของ การบีบอัดสัญญาณวิดีโอ

2.2.4 การแปลงและอินเวอร์สการแปลง (Transform and Inverse Transform)

มาตรฐาน H.263 ใช้การแปลงแบบ DCT ขนาดบล็อก 8x8 จุดภาพ แต่สำหรับมาตรฐาน H.264 ใช้การแปลงแบบจำนวนเต็ม (Integer Transform) ขนาดบล็อก 4x4 จุดภาพ การแปลงแบบจำนวนเต็มใช้พื้นฐานจากการแปลงแบบ DCT ในเมตริกการแปลงจะมีเพียงจำนวนเต็มที่อยู่ในช่วง -2 ถึง 2 ซึ่งทำให้ในการคำนวณจะใช้เพียง การเลื่อน (Shift) การเพิ่ม (Add) และการแยก (Subtract) ซึ่งความผิดพลาดที่เกิดจากการอินเวอร์สการแปลงจะสามารถหลีกเลี่ยง และแก้ไขได้ ไม่เหมือนกับมาตรฐานก่อนหน้าที่มีปัญหา โดยถ้าบล็อกนั้นเข้ารหัสแบบภายในเฟรม ค่าสัมประสิทธิ์กระแสตรง DC (Direct Coefficient) ของบล็อกส่วนประกอบความส่องสว่างจะถูกนำมาจัดเรียงและทำการส่งไปก่อน แล้วจึงทำการส่งข้อมูลของบล็อกที่เหลือแสดงดังรูปที่ 2.7



รูปที่ 2.7 การจัดเรียงค่าสัมประสิทธิ์

อินพุท X ขนาด 4×4 การแปลงสามารถคำนวณได้จากสมการที่ (2.2)

$$Y = AXA^T = \begin{bmatrix} a & a & a & a \\ b & c & -c & -b \\ a & -a & -a & a \\ c & -b & b & -c \end{bmatrix} X \begin{bmatrix} a & b & a & c \\ a & c & -a & -b \\ a & -c & -a & b \\ a & -b & a & -c \end{bmatrix} \quad (2.2)$$

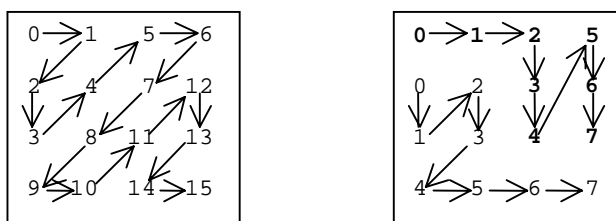
$$a = \frac{1}{2}$$

$$\text{เมื่อ } b = \frac{1}{\sqrt{2}} \cos\left(\frac{\pi}{8}\right)$$

$$c = \frac{1}{\sqrt{2}} \cos\left(\frac{3\pi}{8}\right)$$

2.2.5 การควอนไทซ์และการสแกนค่าสัมประสิทธิ์ที่ได้

การควอนไทซ์เป็นขั้นตอนของการสูญเสียข้อมูลบางส่วน การสูญเสียข้อมูลขึ้นอยู่กับค่าควอนไทซ์ กล่าวคือถ้าขนาดของการควอนไทซ์มีค่ามากจะสูญเสียข้อมูลไปมาก ขนาดของการควอนไทซ์มีค่าน้อยจะสูญเสียข้อมูลไปน้อย ทั้งนี้การเลือกค่าควอนไทซ์เป็นขั้นตอนที่มีความสำคัญต่อจำนวนบิตที่ทำการเข้ารหัสเฟรมนั้น ๆ อีกด้วย สำหรับมาตรฐาน H.263 ได้กำหนดขั้นของการควอนไทซ์ (quantization step) ไว้ที่ 32 ระดับ แต่ในมาตรฐาน H.264 ได้กำหนดขั้นของการควอนไทซ์ (quantization step) ไว้ที่ 51 ระดับ ค่าสัมประสิทธิ์ของการแปลงแบบจำนวนเต็ม ค่าสัมประสิทธิ์ที่อยู่ทางด้านบนซ้ายมือจะแทนค่าสัมประสิทธิ์กระแสตรง DC (Direct Current Coefficient) มีค่ามากที่สุด และที่เหลือเรียกว่าสัมประสิทธิ์กระแสสลับ AC (Alternate Current Coefficient) และเมื่อผ่านการควอนไทซ์เรียงค่าสัมประสิทธิ์ให้เป็นมิติเดียวโดยเริ่มจากค่าสัมประสิทธิ์ DC ซึ่งในมาตรฐาน H.264 นี้จะมีวิธีเรียงค่าสัมประสิทธิ์อยู่ 2 วิธี คือการกราดซิกแซก (Zigzag scan) เป็นวิธีที่ใช้อยู่โดยทั่วไปทั้งในมาตรฐานก่อนหน้าวิธีที่ 2 คือการกราดแบบคู่ (Double scan) ลักษณะการเรียงค่าสัมประสิทธิ์ทั้ง 2 วิธี แสดงดังรูปที่ 2.8



รูปที่ 2.8 วิธีการกราดค่าสัมประสิทธิ์แบบซิกแซกและแบบคู่

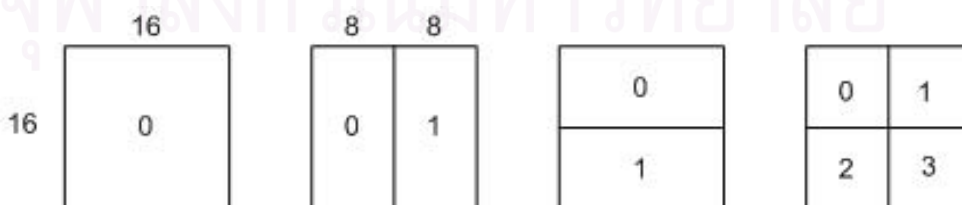
2.2.6 การเข้ารหัสเอ็นโทรปี

เป็นขั้นตอนการแปลงค่าจากสัมประสิทธิ์ทางความถี่ที่ได้จากการควอนไทซ์ให้เป็นเลขไบนารี เพื่อส่งเป็นบิตสตรีมออกไป โดยมาตรฐาน H.264 มีวิธีการเข้ารหัสเอ็นโทรปี 2 วิธี คือ Context-Adaptive Variable Length Codec (CAVLC) [23] และ Context-Based Adaptive Arithmetic Coding (CABAC) [23] โดย CABAC สามารถลดอัตราบิตได้มากกว่า CAVLC ประมาณ 5%-15% [20]

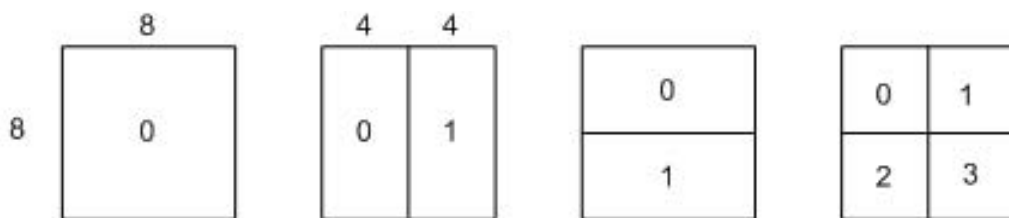
2.2.7 การประมาณการเคลื่อนที่ (Motion Estimation) และการชดเชยการเคลื่อนที่ (Motion compensation)

การประมาณการเคลื่อนที่เป็นการหาความแตกต่างระหว่างจุดภาพที่สอดคล้องกันระหว่างเฟรมสองเฟรมที่ติดกัน โดยปกติเฟรมที่อยู่ติดกันจะมีความสัมพันธ์กันมาก หรือที่เรียกว่า ความซ้ำซ้อนทางเวลาส่งผลให้สามารถเข้ารหัสเฉพาะความแตกต่างแทนที่จะเข้ารหัสทั้งภาพ โดยทั่วไปเราสามารถสมมุติว่าค่าของจุดภาพในภาพที่มีการเคลื่อนที่ของวัตถุได้เลื่อนจากจุดหนึ่งในเฟรมหนึ่งไปยังจุดอื่นในเฟรมถัดไป ถ้าเราสามารถหาตำแหน่งที่เปลี่ยนแปลงหรือที่เรียกว่า เวกเตอร์การเคลื่อนที่ ทำให้สามารถลดความแตกต่างลงไปได้ วิธีนี้เรียกว่าการชดเชยการเคลื่อนที่

โดยทั่วไปขนาดของบล็อกที่ใช้ในการชดเชยการเคลื่อนที่มีขนาด 16×16 หรือ 8×8 จุดภาพ ซึ่งในมาตรฐาน H.264 สำหรับบล็อกที่มีขนาด 16×16 สามารถแบ่งขนาดบล็อกให้มีขนาดเล็กได้ทั้งหมด 4 แบบ คือ 16×16 16×8 8×16 8×8 จุดภาพ แสดงดังรูปที่ 2.9 และสำหรับบล็อกที่มีขนาด 8×8 สามารถแบ่งให้มีขนาดเล็กได้อีก 4 แบบ คือ 8×8 8×4 4×8 และ 4×4 จุดภาพ แสดงดังรูปที่ 2.10

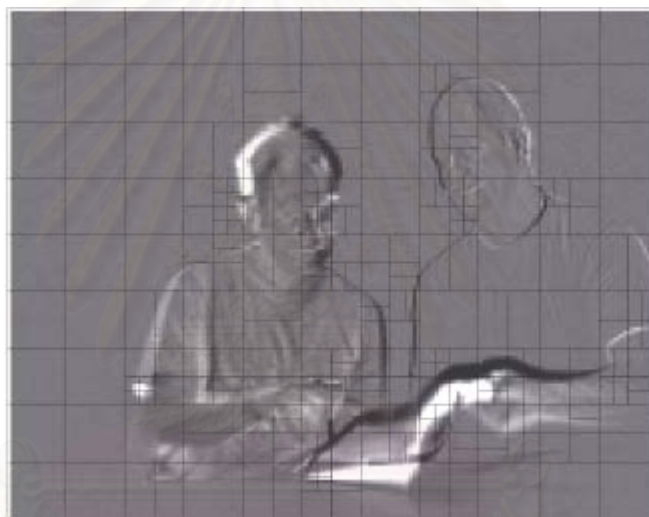


รูปที่ 2.9 การแบ่งบล็อกขนาด 16×16 จุดภาพ



รูปที่ 2.10 การแบ่งบล็อกขนาด 8x8 จุดภาพ

สำหรับการเลือกขนาดของบล็อกโดยทั่วไป บล็อกขนาดใหญ่เหมาะสำหรับพื้นที่ที่มีลักษณะเอกพันธ์ (homogeneous) และบล็อกขนาดเล็กเหมาะสำหรับพื้นที่ที่มีรายละเอียดมาก การแบ่งบล็อกแสดงดังรูปที่ 2.11 [22]



รูปที่ 2.11 แสดงการแบ่งบล็อกภายในภาพ [22]

เมื่อทำการหาบล็อกที่มีจุดภาพใกล้เคียงมากที่สุดระหว่างเฟรมปัจจุบันและเฟรมก่อนหน้า ในการค้นหาบล็อกที่ใกล้เคียงสามารถหาได้ในพื้นที่การค้นหา (Search area) โดยวิธีที่ใช้วัดความใกล้เคียงระหว่างบล็อกทั้งสองนั้นมีหลายวิธี เช่น ค่าความแตกต่างสัมบูรณ์เฉลี่ย (Mean Absolute Difference: MAD) ดังสมการที่ (2.3) ค่าความแตกต่างกำลังสองรวม (Sum Square Difference: SSD) ดังสมการที่ (2.4) ความแตกต่างสัมบูรณ์รวม (Sum Absolute Difference: SAD) แสดงดังรูปที่ 2.12 โดยตำแหน่งที่เปลี่ยนแปลงเรียกว่าเวกเตอร์การเคลื่อนที่

$$MAD(dx, dy) = \frac{1}{N \times N} \sum_{m=x}^{x+N-1} \sum_{n=y}^{y+N-1} |I_k(m, n) - I_{k-1}(m + dx, n + dy)| \quad (2.3)$$

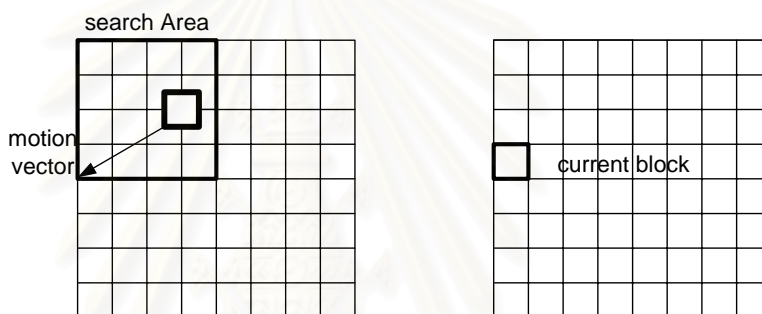
$$SSD(dx, dy) = \frac{1}{N \times N} \sum_{m=x}^{x+N-1} \sum_{n=y}^{y+N-1} (I_k(m, n) - I_k(m + dx, n + dy))^2 \quad (2.4)$$

โดยที่ $N \times N$ คือขนาดของบล็อก

$I_k(m, n)$ คือค่าความส่องสว่างของเฟรม k ที่ตำแหน่ง (m, n)

$MAD(dx, dy)$ คือค่าความแตกต่างสัมบูรณ์เฉลี่ย เมื่อเวกเตอร์การเคลื่อนที่เท่ากับ (m, n)

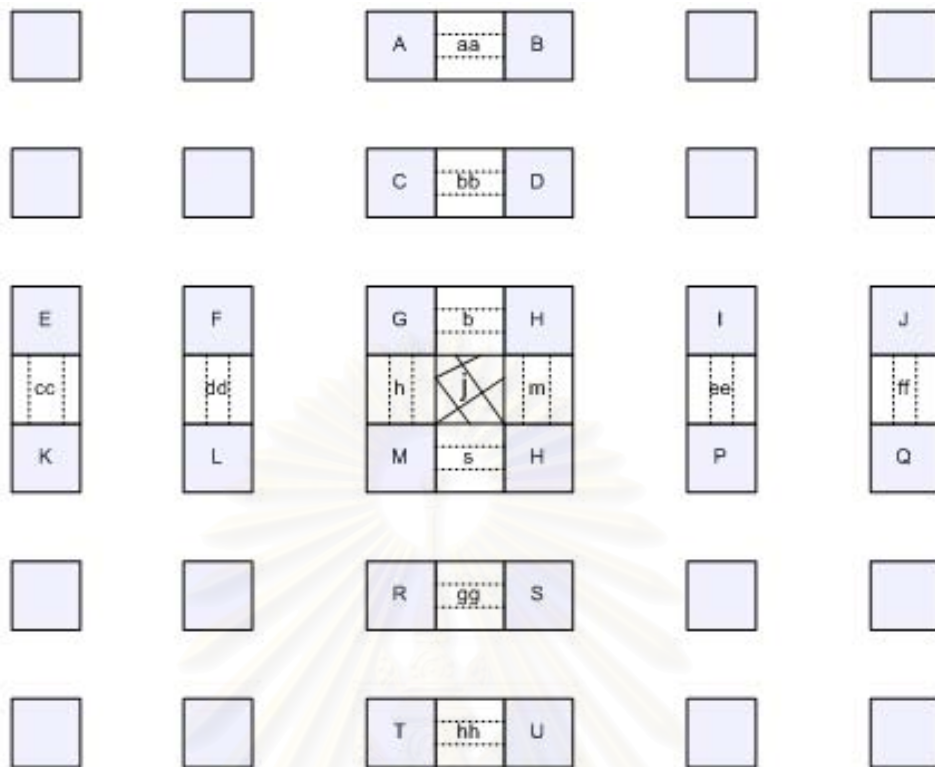
$SSD(dx, dy)$ คือค่าความแตกต่างกำลังสองรวม เมื่อเวกเตอร์การเคลื่อนที่เท่ากับ (m, n)



รูปที่ 2.12 การประมาณการเคลื่อนที่

สำหรับมาตรฐาน H.264 ได้เพิ่มระดับการประมาณการเคลื่อนที่ใหม่จากเดิมแบบครึ่งจุดภาพ (Half-pixel Motion Estimation) เป็น 1/4 และ 1/8 จุดภาพ ซึ่งสามารถทำให้ประมาณการเคลื่อนที่ได้แม่นยำยิ่งขึ้น สำหรับการประมาณการเคลื่อนที่ใหม่จากเดิมแบบครึ่งจุดภาพสามารถแสดงได้ดังรูปที่ 2.13

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 2.13 การหาค่าจุดภาพระหว่าง 2 จุดภาพ

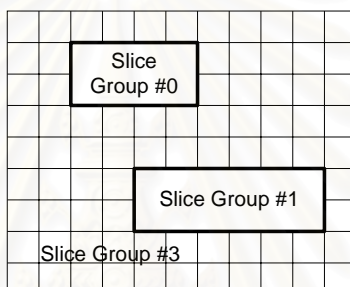
การหาค่าระหว่างจุดภาพสามารถหาได้จากการประมาณในช่วง (Interpolated) โดยใช้ วงจรกรองตอบสนองของอิมพัลส์จำกัด (Finite Impulse Response (FIR) filter) 6 แถบ ซึ่งวงจรกรอง มีค่าถ่วงน้ำหนักคือ $(1/32 \ -5/32 \ 5/8 \ 5/8 \ -5/32 \ 1/32)$ [22] เช่น ต้องการหาจุดภาพ b สามารถ คำนวณได้จากจุดภาพ E F G H I และ J จากสมการที่ (2.5)

$$b = \text{round} ((E-5F+20G+20H-5I+J)/32) \quad (2.5)$$

นอกจากนี้มาตรฐาน H.264 ได้มีการเพิ่มโหมดการทำงานเพื่อการบีบอัดสัญญาณวิดีโอให้ดีขึ้น [20] ดังนี้

1. การคำนวณเวกเตอร์การเคลื่อนที่แบบไม่จำกัด (Unrestricted Motion Vector Mode) นั่นคือถ้าเวกเตอร์การเคลื่อนไหวซึ่งออกนอกพื้นที่ภาพ จะใช้วิธีการประมาณนอกช่วง(extrapolation) ซึ่งเทคนิคนี้อยู่ในโหมดพิเศษในมาตรฐาน H.263
2. การทำนายด้วยค่าถ่วงน้ำหนัก (Weighted prediction) การชดเชยการเคลื่อนที่ที่สามารถถ่วงน้ำหนักบริเวณที่สนใจ

3. การถอดคู่ควบลำดับเฟรมอ้างอิงจากลำดับการแสดงผลภาพ (Decoupling of reference)
4. สามารถอ้างอิงการประมาณการเคลื่อนที่และการชดเชยการเคลื่อนที่สูงสุด 5 เฟรม ทำให้สามารถประมาณการเคลื่อนที่ได้แม่นยำมากขึ้น แต่ต้องคำนึงถึงหน่วยความจำที่ใช้เก็บรายละเอียดของเฟรมก่อนหน้า โดยเทคนิคใช้กับมาตรฐาน H.263++
5. สามารถจัดแบ่งกลุ่มของมาโครบล็อก โดยเรียกว่าการเรียงลำดับมาโครบล็อกแบบยืดหยุ่น (Flexible Macroblock Ordering) ประโยชน์ของการเรียงลำดับมาโครบล็อกแบบยืดหยุ่นนี้สำหรับการแบ่งบริเวณที่สนใจ (Region of Interest) แสดงดังรูปที่ 2.14



รูปที่ 2.14 การจัดกลุ่มมาโครบล็อกแบบยืดหยุ่น

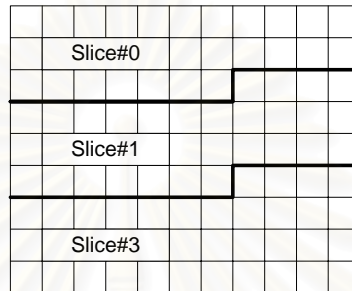
6. สามารถเข้ากับโครงข่ายต่าง ๆ ได้ง่าย (Network friendly)
7. สามารถแบ่งข้อมูลออกเป็นส่วน ๆ (Data Partitioning) โดยถ้าข้อมูลส่วนนั้นมีความสำคัญจะทำการส่งข้อมูลส่วนนั้นก่อน โดยเทคนิคใช้กับมาตรฐาน H.263++

2.2.8 รูปแบบการจัดเรียงสัญญาณ

ในมาตรฐาน H.264 ได้กำหนดรูปแบบการจัดเรียงข้อมูลในสัญญาณที่ถูกเข้ารหัสตามลำดับชั้นต่าง ๆ ทั้งนี้การจัดเรียงข้อมูลรวมทั้งข้อมูลในส่วนหัวของสัญญาณต้องเป็นไปตามมาตรฐาน เพื่อไม่ให้ตัวถอดรหัสเกิดความสับสน โดยจะแบ่งเป็นลำดับชั้นต่าง ๆ ได้ดังนี้

1. ชั้นกลุ่มภาพ (Group Of Picture Layer) ประกอบด้วยจำนวนเฟรมหลายๆเฟรมอยู่ในกลุ่มเดียวกัน และแต่ละกลุ่มภาพจะเริ่มต้นด้วยภาพแบบ I มีเพียงเฟรมเดียวเท่านั้น

2. ชั้นรูปภาพ (Picture layer) จะประกอบไปด้วยจุดภาพขนาดเท่ากับสัญญาณวิดีโอทัศน์ โดยทั่ว ๆ ไปใน 1 เฟรมจะประกอบด้วย 2 ส่วน คือบรรทัดคู่ และบรรทัดคี่
3. ชั้นสไลด์และกลุ่มสไลด์ (Slices and Slice Groups) โดยแต่ละสไลด์จะประกอบด้วยลำดับของมาโครบล็อกซึ่งได้มาจากการเรียงแบบราสเตอร์ (raster scan) การแบ่งกลุ่มสไลด์แสดงดังรูปที่ 2.15



รูปที่ 2.15 การเรียงลำดับมาโครบล็อกแบบปกติ

4. ชั้นมาโครบล็อก (Macroblock layer) แต่ละมาโครบล็อกจะประกอบไปด้วยจุดภาพขนาด 16×16 จุดภาพ
5. ชั้นบล็อก (Block layer) ประกอบไปด้วยจุดภาพขนาด 8×8 จุดภาพ

2.2.9 รูปแบบสัญญาณวิดีโอทัศน์ที่ใช้

รูปแบบที่ใช้ในมาตรฐาน H.263 คือ CIF (Common Intermediate Format) และ QCIF (Quarter CIF) ซึ่งมีความละเอียดลดลงครึ่งหนึ่ง แต่ละรูปแบบประกอบด้วยส่วนประกอบความเข้มสี และความส่องสว่าง โดยรูปแบบ CIF จะมีส่วนประกอบความส่องสว่าง Y มีจำนวนจุดภาพในแนวนอนเป็น 352 จุดภาพ และในแนวตั้ง 288 จุดภาพ ส่วนในรูปแบบ QCIF จะมีความละเอียดลดลงครึ่งหนึ่ง นั่นคือจำนวนจุดภาพแนวนอนเป็น 176 จุดภาพ และในแนวตั้ง 144 จุดภาพ ในส่วนประกอบความเข้มสีของรูปแบบ CIF และ QCIF (Cb และ Cr) จะมีความละเอียดน้อยกว่าส่วนประกอบความส่องสว่าง เนื่องจากตาคนไวต่อความส่องสว่างแต่ไม่ไวต่อความเข้มสี จึงไม่จำเป็นต้องใช้ข้อมูลของความเข้มสีละเอียดมากนัก

2.3 การควบคุมอัตราสำหรับการเข้ารหัสสัญญาณวิดีโอ H.264 [24]

การควบคุมอัตราสำหรับ H.264 แบ่งออกได้ 3 ระดับ คือระดับกลุ่มภาพ (GOP level) ระดับเฟรม (Frame level) ระดับพื้นฐานย่อย (Basic unit level) สามารถแสดงดังรูปที่ 2.16 โดยที่ในระดับกลุ่มภาพทำการคำนวณจำนวนบิตสำหรับการเข้ารหัสกลุ่มภาพนั้น ๆ จากนั้นในระดับเฟรมภาพ (Frame level) ทำการคำนวณหาจำนวนบิตเป้าหมายที่ทำการเข้ารหัสในแต่ละเฟรม ส่วนในระดับหน่วยพื้นฐานย่อย (Basic unit level) จะทำการหาจำนวนบิตเป้าหมายในหน่วยพื้นฐานย่อยนั้น ๆ เมื่อได้จำนวนบิตเป้าหมายในระดับหน่วยพื้นฐานย่อยแล้วจะทำการหาค่าควอนไทซ์พารามิเตอร์โดยใช้โมเดลกำลังสอง ซึ่งการหาค่าควอนไทซ์พารามิเตอร์นี้ใช้ค่าความแตกต่างสัมบูรณ์ (Mean Absolute Difference) ด้วย จากนั้นทำการตรวจสอบเงื่อนไขเพื่อกำหนดค่าควอนไทซ์พารามิเตอร์ เมื่อเข้ารหัสครบทุกหน่วยพื้นฐานย่อยแล้วทำการรวมจำนวนบิตทั้งหมดภายในเฟรม แล้วทำการคำนวณหาจำนวนบิตสะสมคงเหลือในบัฟเฟอร์ โดยบล็อกไดอะแกรมการควบคุมอัตราสำหรับการเข้ารหัสสัญญาณวิดีโอ H.264 สามารถแสดงได้ดังรูปที่ 2.16

2.3.1 การควบคุมอัตราระดับกลุ่มภาพ

ในระดับนี้คำนวณหาจำนวนบิตสำหรับการเข้ารหัสภาพภายในกลุ่มภาพ ในกรณีที่เฟรมแรกของกลุ่มภาพ จำนวนบิตทั้งหมดคำนวณดังสมการที่ (2.6)

$$T_r(n_{i,0}) = \frac{u(n_{i,1})}{F_r} * N_{gop} - \left(\frac{B_s}{8} - B_c(n_{i-1, N_{gop}}) \right) \quad (2.6)$$

โดยที่ $n_{i,j}$ ($i=1,2,\dots, j=1,2,\dots, N_{gop}$) คือลำดับภาพที่ j ในกลุ่มภาพที่ i

T_r คือจำนวนบิตสำหรับกลุ่มภาพ

N_{gop} คือจำนวนเฟรมในแต่ละกลุ่มภาพ (GOP)

F_r คืออัตราเฟรม

B_s คือขนาดบัฟเฟอร์เริ่มต้นโดยเท่ากับสองเท่าของอัตรา

$B_c(n_{i,j})$ คือจำนวนบิตที่สะสมอยู่ในบัฟเฟอร์หลังจากเข้ารหัสแต่ละเฟรม

$u(n_{i,j})$ คืออัตราในช่องสัญญาณ ในเฟรมที่ j ในกลุ่มภาพที่ i

โดยการคำนวณจำนวนบิตสะสมในบัพเฟออร์สามารถคำนวณโดยใช้โมเดลทราฟฟิกการไหลของไหล โดยจำนวนบิตสะสมในบัพเฟออร์หลังจากเข้ารหัสเฟรม j ($B_c(n_{i,j+1})$) เลือกค่าน้อยที่สุดระหว่างค่าบัพเฟออร์เริ่มต้นกับค่ามากที่สุดระหว่างค่าศูนย์กับจำนวนบิตสะสมในบัพเฟออร์ของเฟรมก่อนหน้ากับผลต่างระหว่างจำนวนบิตที่ใช้ในการเข้ารหัสเฟรมก่อนหน้ากับจำนวนบิตที่สามารถออกสู่ช่องสัญญาณได้ แสดงในสมการที่ (2.7)

$$B_c(n_{i,j+1}) = \min\left\{\max\left\{0, B_c(n_{i,j}) + A(n_{i,j}) - \frac{u(n_{i,j})}{F_r}\right\}, B_s\right\}$$

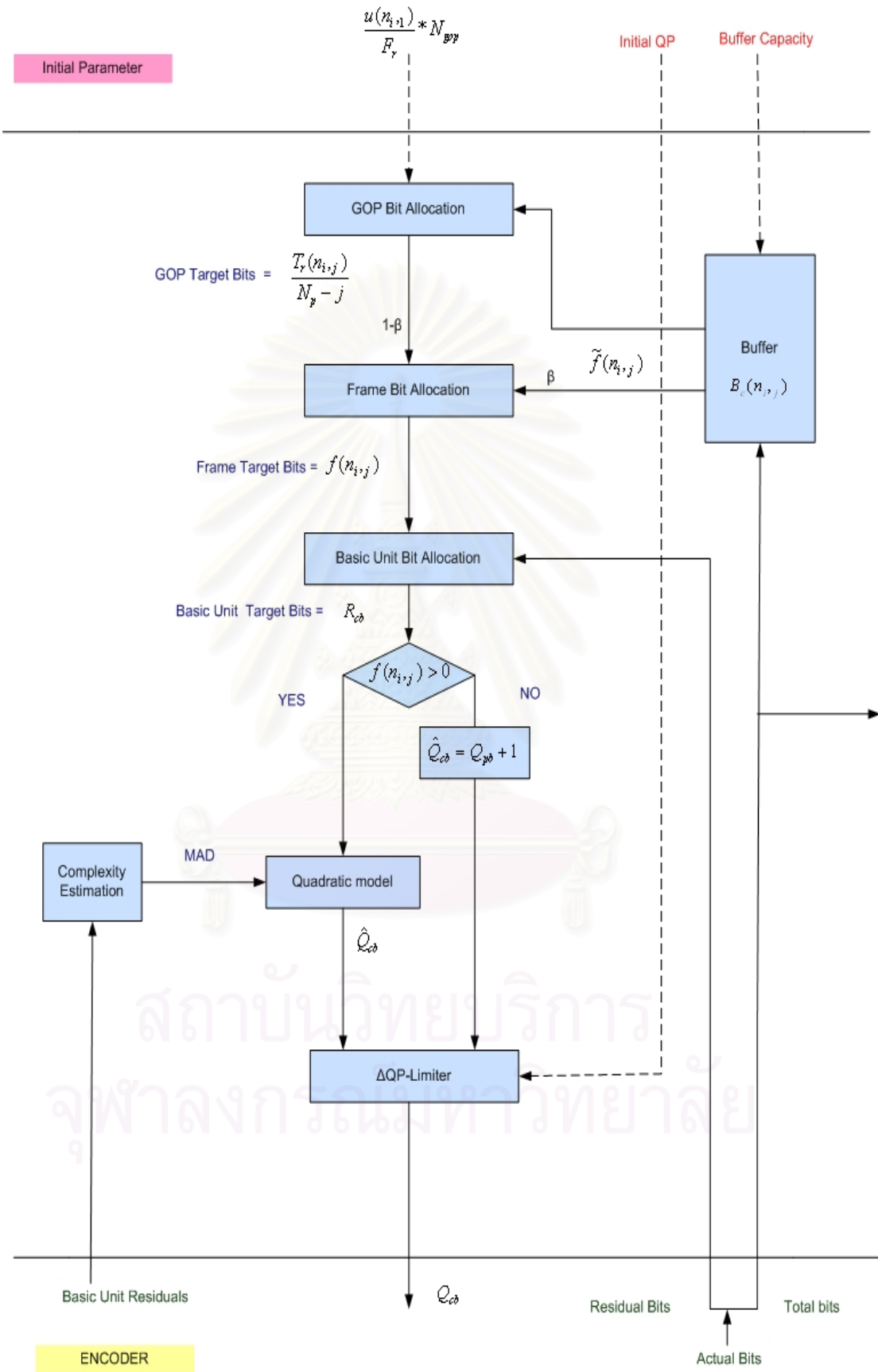
$$B_c(n_{i,1}) = \frac{B_s}{8} \tag{2.7}$$

$$B_c(n_{i+1,0}) = B_c(n_{i,N_{gap}})$$

เมื่อ $A(n_{i,j})$ คือจำนวนบิตที่ใช้ในการเข้ารหัสเฟรมที่ j ในกลุ่มภาพที่ i

ดังนั้นถ้าทำการเข้ารหัสแต่ละเฟรมเสร็จต้องทำการคำนวณจำนวนบิตคงเหลือสามารถคำนวณดังแสดงในสมการที่ (2.8)

$$T_r(n_{i,j}) = T_r(n_{i,j-1}) - A(n_{i,j-1}) \tag{2.8}$$



รูปที่ 2.16 บล็อกไดอะแกรมการควบคุมอัตราสำหรับมาตรฐาน H.264

2.3.2 การควบคุมอัตราระดับเฟรม

ในระดับนี้แบ่งการทำงานได้ เป็น 2 ขั้นตอน

1. Pre-Encoding Stage ทำการคำนวณหาจำนวนบิตเป้าหมายเพื่อเข้ารหัสในแต่ละเฟรม แต่สำหรับกรณีที่เข้ารหัสเพียง 2 ระดับ คือระดับกลุ่มภาพ และระดับเฟรม ในขั้นตอนนี้จะคำนวณค่าควอนไทซ์พารามิเตอร์ด้วย
2. Post-Encoding Stage ทำการปรับค่าพารามิเตอร์ต่าง ๆ หลังการเข้ารหัสแต่ละเฟรม

2.3.2.1 ส่วน Pre-Encoding Stage

ในขั้นตอนนี้คำนวณหาจำนวนบิตเป้าหมายในแต่ละเฟรม และการคำนวณค่าควอนไทซ์พารามิเตอร์โดยมี 2 ขั้นตอนดังนี้

ขั้นตอนที่ 1 กำหนดบิตเป้าหมาย (Target bits) สำหรับเฟรมภาพ โดยบิตเป้าหมายสามารถคำนวณจาก 2 ส่วน ส่วนที่มาจากระดับกลุ่มภาพและส่วนที่มาจากระดับบัฟเฟอร์ขั้นตอนนี้แบ่งได้ 2 ขั้นตอนย่อยดังนี้

ขั้นตอนที่ 1.1 การควบคุมมหรรศน์ (Macroscopic Control)

การทำนายระดับบัฟเฟอร์เป้าหมาย (Target Buffer level) เพื่อนำไปใช้ในการคำนวณบิตเป้าหมายแสดงในสมการที่ (2.9)

$$T_{bl}(n_{i,j+1}) = T_{bl}(n_{i,j}) - \frac{B_c(n_{i,2}) - B_s / 8}{N_p - 1} \quad (2.9)$$

เมื่อ $T_{bl}(n_{i,j})$ คือระดับบัฟเฟอร์เป้าหมายในเฟรมที่ j ในกลุ่มภาพที่ i

N_p คือจำนวนเฟรมภาพ P ในกลุ่มภาพ

ขั้นตอนที่ 1.2 การควบคุมจุลทรรศน์ (Microscopic control)

จากที่กล่าวข้างต้นการคำนวณบิตเป้าหมาย $f(n_{i,j})$ ประกอบด้วย 2 ส่วน ส่วนที่ใช้ระดับบัพเฟอร์ ($\tilde{f}(n_{i,j})$) และส่วนที่มาจากระดับกลุ่มภาพ

ส่วนที่มาจากระดับบัพเฟอร์สามารถคำนวณจากจำนวนบิตในบัพเฟอร์ ระดับบัพเฟอร์เป้าหมาย อัตราเฟรมและอัตราในช่องสัญญาณดังแสดงในสมการที่ (2.10)

$$\tilde{f}(n_{i,j}) = \frac{u(n_{i,j})}{F_r} + \gamma(Tbl(n_{i,j}) - B_c(n_{i,j})) \quad (2.10)$$

เมื่อ γ คือค่าคงที่ ในที่นี้กำหนดให้มีค่าเท่ากับ 0.75

ส่วนที่มาจากระดับกลุ่มภาพหาได้จากจำนวนบิตคงเหลือสำหรับเข้ารหัสเฟรมในกลุ่มภาพหารด้วยจำนวน P เฟรมคงเหลือแล้วทำการถ่วงน้ำหนักบิตเป้าหมายเพื่อให้ค่ามีความถูกต้องมากขึ้นดังแสดงในสมการที่ (2.11)

$$f(n_{i,j}) = \beta * \frac{T_r(n_{i,j})}{N_p - j} + (1 - \beta) * \tilde{f}(n_{i,j}) \quad (2.11)$$

เมื่อ β คือค่าถ่วงน้ำหนักมีค่าตั้งแต่ 0 ถึง 1 ในที่นี้กำหนดให้มีค่าเท่ากับ 0.5

ขั้นตอนที่ 2 คำนวณค่าควอนไทซ์พารามิเตอร์โดยใช้โมเดลกำลังสอง [19] [25] [26] ดังแสดงในสมการที่ (2.12)

$$R_i = b_1 \times Q_i^{-1} + b_2 \times Q_i^{-2} \quad (2.12)$$

เมื่อ R_i คือจำนวนบิตที่จะเข้ารหัสเฟรม i Q_i คือค่าควอนไทซ์สำหรับเฟรม i b_1 และ b_2 คือค่าสัมประสิทธิ์อันดับที่ 1 และอันดับที่ 2 ตามลำดับ

การหาค่าควอนไทล์สำหรับโมเดลกำลังสอง ต้องพิจารณาอีก 2 ปัจจัยคือ

1. โมเดลกำลังสองพัฒนาจากสมมุติฐานที่ว่าเฟรมแต่ละเฟรมที่อยู่ติดกันสามารถเข้ารหัสด้วยความซับซ้อนที่คล้ายกันโดยไม่คำนึงถึงความซับซ้อนของแต่ละเฟรมจึงต้องมีดัชนี (index) ที่ใช้บอกความซับซ้อน เช่น ค่าความแตกต่างสัมบูรณ์เฉลี่ย
2. โมเดลนี้ยังได้รวมถึงข้อมูลส่วนหัว (overhead data) และเวกเตอร์การเคลื่อนที่ซึ่งต้องลบออกจากจำนวนบิตที่จะเข้ารหัสเฟรม เพื่อให้ได้จำนวนบิตที่เข้ารหัสเฟรมนั้นอย่างแท้จริง ดังนั้นการหาค่าควอนไทล์โดยพิจารณา 2 ปัจจัยข้างต้นสามารถคำนวณได้จากสมการที่ (2.13)

$$\frac{R_i - H_i}{M_i} = b_1 \times Q_i^{-1} + b_2 \times Q_i^{-2} \quad (2.13)$$

เมื่อ M_i คือค่าความแตกต่างสัมบูรณ์เฉลี่ยสำหรับเฟรม i โดยค่าความแตกต่างสัมบูรณ์เฉลี่ยหาได้จากโมเดลเชิงเส้นโดยค่าความแตกต่างสัมบูรณ์เฉลี่ยสำหรับปัจจุบันโดยใช้โมเดลเชิงเส้น โดยคำนวณเทียบจากตำแหน่งเดียวกันของเฟรมก่อนหน้า จากสมการที่ (2.14)

$$MAD_f = a_1 * MAD_{f-1} + a_2 \quad (2.14)$$

เมื่อ MAD_f คือค่าความแตกต่างสัมบูรณ์เฉลี่ยของเฟรมปัจจุบัน

MAD_{f-1} คือค่าความแตกต่างสัมบูรณ์เฉลี่ยของเฟรมก่อนหน้า

a_1, a_2 คือค่าสัมประสิทธิ์ โดยเริ่มต้นกำหนดให้มีค่าเท่ากับ 1 และ 0 ตามลำดับ

H_i คือข้อมูลส่วนหัวและเวกเตอร์การเคลื่อนที่สำหรับเฟรม i

ดังนั้นจากสมการที่ (2.11) สามารถหาคำนวนค่าควอนไทล์ดังแสดงในสมการที่ (2.15)

$$\frac{f(n_{i,j}) - H(n_{i,j})}{M(n_{i,j})} = a_1 \times \hat{Q}_{pc}^{-1} + a_2 \times \hat{Q}_{pc}^{-2} \quad (2.15)$$

เมื่อทำการคำนวณหาค่าควอนไทล์ที่จะทำการเข้ารหัสเฟรมปัจจุบัน (\hat{Q}_{pc}) เพื่อให้คุณภาพสัญญาณวีดิทัศน์ราบเรียบต่อเนื่องกันมากขึ้น จึงทำการปรับค่าควอนไทล์พารามิเตอร์

ไม่ให้ค่าควอนไทล์ที่จะทำการเข้ารหัสเฟรมปัจจุบันกับค่าควอนไทล์ที่ทำการเข้ารหัสเฟรมก่อนหน้าต่างกันเกิน 2 ระดับ โดยแสดงในสมการที่ (2.16)

$$\tilde{Q}_{pc} = \min\{Q_{pp} + 2, \max\{Q_{pp} - 2, \hat{Q}_{pc}\}\} \quad (2.16)$$

เมื่อ Q_{pp} คือค่าควอนไทล์ที่ทำการเข้ารหัสเฟรมก่อนหน้า

ปรับค่าควอนไทล์ที่ได้ต้องไม่เกินค่าต่ำสุด และค่าสูงสุดของขั้นการควอนไทล์ ดังแสดงในสมการที่ (2.17)

$$Q_{pc} = \min\{51, \max\{\tilde{Q}_{pc}, 1\}\} \quad (2.17)$$

เมื่อได้ค่าควอนไทล์แล้วจึงนำไปใช้เพื่อหาโหมด (MODE) ที่เหมาะสมสำหรับแต่ละมาโครบล็อก [27] ดังแสดงในสมการที่ (2.18)

$$D(s, c, MODE | QP) + \lambda_{mode} R(s, c, MODE | QP) \quad (2.18)$$

เมื่อ $\lambda_{mode} = 0.85 \times 2^{QP/3}$ สำหรับ I, P เฟรม

เมื่อ *MODE* คือ INTRA SKIP INTER INTER+4V INTER+Q(-2) INTER+Q(-1) INTER+Q(1) และ INTER+Q(2) โดยที่ INTER+Q(c) มาโครบล็อกที่อยู่ติดกันค่าควอนไทล์เปลี่ยนแปลง ไม่เกิน c ระดับ

โดยค่า $D(s, c | MODE)$ วัดโดยใช้ค่าความแตกต่างกำลังสองรวมระหว่างมาโครบล็อกที่ถูกสร้างขึ้น (s) กับมาโครบล็อกต้นแบบ (c) หาได้จากสมการที่ (2.19)

$$D(s, c, MODE | QP) = \sum_{x=1, y=1}^{16,16} (s_y[x, y] - c_y[x, y, MODE | QP])^2 + \sum_{x=1, y=1}^{8,8} (s_U[x, y] - c_U[x, y, MODE | QP])^2 + \sum_{x=1, y=1}^{8,8} (s_V[x, y] - c_V[x, y, MODE | QP])^2, \quad (2.19)$$

โดย $c_Y[x, y, MODE | QP]$ คือความส่องสว่างของมาโครบล็อกที่ถูกสร้างขึ้น $s_Y[x, y]$ คือความส่องสว่างของมาโครบล็อกต้นแบบ c_U c_V s_U และ s_V คือสัญญาณสีของมาโครบล็อก ส่วน $R(s, c | MODE)$ คือจำนวนบิตที่ได้จากการเลือกโหมดและค่าควอนไทซ์

2.3.2.2 ส่วน Post-Encoding Stage

ทำการปรับค่าพารามิเตอร์ต่างๆ เช่น a_1 และ a_2 ในโมเดลเชิงเส้น b_1 และ b_2 ในโมเดลกำลังสองสำหรับหาค่าควอนไทซ์พารามิเตอร์

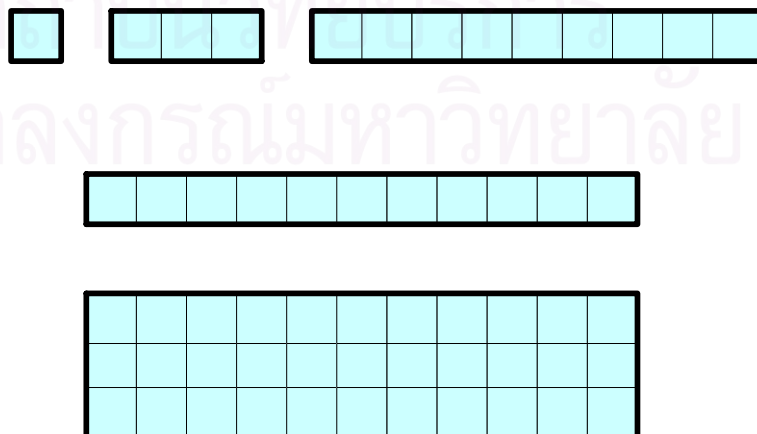
2.3.3 การควบคุมอัตราระดับหน่วยพื้นฐานย่อย

ในส่วนนี้จะกล่าวถึงนิยามของหน่วยพื้นฐานย่อยและการควบคุมอัตราระดับหน่วยพื้นฐานย่อย

2.3.3.1 นิยามของหน่วยพื้นฐานย่อย (Basic unit)

ในหนึ่งเฟรมประกอบด้วยมาโครบล็อกหลายๆ มาโครบล็อก เช่น QCIF มีทั้งหมด 99 มาโครบล็อก โดยขนาดหนึ่งหน่วยพื้นฐานย่อย (N_{mbpic}) คือกลุ่มของมาโครบล็อก โดยมีขนาดตั้งแต่ 1 3 9 11 และ 33 แสดงดังรูปที่ 2.17 ดังนั้นจำนวนหน่วยพื้นฐานย่อยในหนึ่งเฟรม N_{unit} สามารถคำนวณได้จากสมการที่ (2.20)

$$N_{unit} = \frac{N_{mbpic}}{N_{mbunit}} \quad (2.20)$$



รูปที่ 2.17 ขนาดหน่วยพื้นฐานย่อย

2.3.3.2 การควบคุมอัตราของหน่วยพื้นฐานย่อย

จุดมุ่งหมายของการควบคุมอัตราของหน่วยพื้นฐานย่อยคือหาค่าควอนไทล์ที่เหมาะสม โดยที่ผลรวมของจำนวนบิตที่ใช้ในแต่ละเฟรมต้องใกล้เคียงกับจำนวนบิตเป้าหมาย $f(n_{i,j})$ ในแต่ละเฟรมมากที่สุด โดยมี 6 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 คำนวณจำนวนบิตที่ใช้สำหรับหน่วยพื้นฐานย่อยปัจจุบัน $R(n_{i,j,k})$ เมื่อ k คือหน่วยพื้นฐานย่อยปัจจุบันที่จะทำการเข้ารหัส โดย k มีค่าตั้งแต่ 1 ถึง N_{mbunit} โดยขั้นตอนนี้ประกอบด้วย 3 ขั้นตอนย่อย

ขั้นตอนที่ 1.1 คำนวณจำนวนบิตเป้าหมายสำหรับหน่วยพื้นฐานย่อยปัจจุบันเมื่อ $f_{rb}(n_{i,j,k})$ คือจำนวนบิตคงเหลือสำหรับเข้ารหัสหน่วยพื้นฐานย่อยตั้งแต่ k ถึง N_{mbunit} และ N_{ub} คือจำนวนหน่วยพื้นฐานย่อยคงเหลือ โดยจำนวนบิตเป้าหมายคำนวณจาก f_{rb}/N_{ub}

ขั้นตอนที่ 1.2 คำนวณจำนวนบิตสำหรับส่วนหัวสำหรับหน่วยพื้นฐานย่อยปัจจุบัน (m_{hdr}) สามารถคำนวณได้จากสมการที่ (2.21)

$$\begin{aligned} \tilde{m}_{hdr,k} &= \left(\tilde{m}_{hdr,k-1} \times \left(1 - \frac{1}{k}\right) \right) + \frac{\hat{m}_{hdr,k}}{k} \\ m_{hdr} &= \left(\tilde{m}_{hdr,k} \times \frac{k}{N_{unit}} \right) + \left(m_{hdr,1} \times \left(1 - \frac{k}{N_{unit}}\right) \right) \end{aligned} \quad (2.21)$$

เมื่อ $\hat{m}_{hdr,k}$ คือจำนวนบิตของส่วนหัวแท้จริง สำหรับหน่วยพื้นฐานย่อยที่ k ในเฟรมปัจจุบัน $m_{hdr,1}$ คือจำนวนบิตของส่วนหัวที่มาจากการประมาณเฟรมก่อนหน้า

ขั้นตอนที่ 1.3 คำนวณจำนวนบิตที่ใช้สำหรับหน่วยพื้นฐานย่อยปัจจุบัน R_{cb} จากสมการที่ (2.22)

$$R_{cb} = \frac{f_{rb}}{N_{ub}} - m_{hdr} \quad (2.22)$$

ขั้นตอนที่ 2 ทำนายค่าความแตกต่างสัมบูรณ์เฉลี่ยสำหรับหน่วยพื้นฐานย่อยปัจจุบัน โดยใช้โมเดลเชิงเส้น โดยคำนวณเทียบจากตำแหน่งเดียวกันของเฟรมก่อนหน้าดังสมการที่ (2.23)

$$MAD_{cb} = a_1 * MAD_{cb-1} + a_2 \quad (2.23)$$

เมื่อ MAD_{cb} คือค่าความแตกต่างสัมบูรณ์เฉลี่ยของหน่วยพื้นฐานปัจจุบัน

MAD_{cb-1} คือค่าความแตกต่างสัมบูรณ์เฉลี่ยของหน่วยพื้นฐานก่อนหน้า

a_1, a_2 คือค่าสัมประสิทธิ์ โดยเริ่มต้นกำหนดให้มีค่าเท่ากับ 1 และ 0 ตามลำดับ

ขั้นตอนที่ 3 คำนวณค่าควอนไทซ์พารามิเตอร์โดยใช้โมเดลกำลังสองเหมือนในระดับกลุ่มเฟรม โดยพิจารณาได้ 3 กรณี

กรณีที่ 1 ค่าควอนไทซ์พารามิเตอร์สำหรับหน่วยพื้นฐานแรกในเฟรมสามารถหาได้จากสมการที่ (2.24)

$$Q_{cb} = Q_{apf} \quad (2.24)$$

เมื่อ Q_{apf} คือค่าเฉลี่ยค่าควอนไทซ์พารามิเตอร์ทั้งเฟรมของเฟรมก่อนหน้า

กรณีที่ 2 ถ้าจำนวนบิตเป้าหมาย (f_{rb}) มีค่าน้อยกว่า 0 ทำการเพิ่มค่าระดับควอนไทซ์พารามิเตอร์เพื่อไม่ให้ผลรวมจำนวนบิตที่เข้ารหัสเกินกว่าจำนวนบิตเป้าหมายมาก สามารถหาได้จากสมการที่ (2.25)

$$\hat{Q}_{cb} = Q_{pb} + DQuant \quad (2.25)$$

เมื่อ Q_{pb} คือค่าควอนไทซ์พารามิเตอร์ของหน่วยพื้นฐานย่อยก่อนหน้า $DQuant$ มีค่าเท่ากับ 1

ปรับค่าควอนไทซ์พารามิเตอร์เพื่อให้คุณภาพสัญญาณวีดิทัศน์ราบเรียบต่อเนื่องกันมากขึ้น และปรับค่าควอนไทซ์ที่ได้ต้องไม่เกินค่าต่ำสุด และค่าสูงสุดของขั้นการควอนไทซ์ ดังสมการที่ (2.26)

$$Q_{cb} = \max\{1, Q_{apf} - \Delta, \min\{51, Q_{apf} + \Delta, \hat{Q}_{cb}\}\} \quad (2.26)$$

กรณีที่ 3 ถ้าจำนวนบิตเป้าหมาย (f_{rb}) มีค่ามากกว่า 0 คำนวณค่าควอนไทล์พารามิเตอร์ตามโมเดลกำลังสองดังสมการที่ (2.27)

$$\frac{R_{cb} - m_{hdr}}{MAD_{cb}} = b_1 \times \hat{Q}_{cb}^{-1} + b_2 \times \hat{Q}_{cb}^{-2} \quad (2.27)$$

เพื่อให้คุณภาพสัญญาณวีดิทัศน์ราบเรียบต่อเนื่องกันมากขึ้นจึงทำการปรับค่าควอนไทล์พารามิเตอร์ดังสมการที่ (2.28)

$$\tilde{Q}_{cb} = \max\{Q_{pb} - DQuant, \min\{\hat{Q}_{cb}, Q_{pb} + DQuant\}\} \quad (2.28)$$

ปรับค่าควอนไทล์ที่ได้ต้องไม่เกินค่าต่ำสุดและค่าสูงสุดของชั้นการควอนไทล์ดังสมการที่ (2.29)

$$Q_{cb} = \max\{1, Q_{apf} - \Delta, \min\{51, Q_{apf} + \Delta, \hat{Q}_{cb}\}\} \quad (2.29)$$

ขั้นตอนที่ 4 เลือกโหมดสำหรับแต่ละหน่วยพื้นฐานย่อย

ขั้นตอนที่ 5 ปรับค่าจำนวนบิตคงเหลือ และจำนวนหน่วยพื้นฐานย่อย

ขั้นตอนที่ 6 ปรับค่า Q_{apf} หลังจากเข้ารหัสครบทุกหน่วยพื้นฐานย่อยในเฟรม

2.4 การวัดคุณภาพของภาพ (Quality Measurement)

ในการวัดคุณภาพของภาพที่สร้างขึ้นจากกระบวนการชดเชยการเคลื่อนที่ ด้วยวิธีประมาณการเคลื่อนที่ เป็นวิธีตรวจสอบว่า ภาพที่ได้มีความใกล้เคียงกับภาพที่ต้องการประมาณมากน้อยเพียงใด ค่าที่ใช้วัดคุณภาพนิยมใช้กันอยู่ 2 ชนิด คือค่าความผิดพลาดกำลังสองเฉลี่ย และค่าอัตราส่วนสัญญาณยอดต่อสัญญาณรบกวน

2.4.1 ค่าความผิดพลาดกำลังสองเฉลี่ย (Mean Square Error: MSE)

เป็นการวัดคุณภาพแบบ Objective สามารถคำนวณดังสมการที่ (2.30)

$$MSE = \frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N (X(m,n) - X^R(m,n))^2 \quad (2.30)$$

โดยที่ $M \times N$ คือขนาดของเฟรม

$X(m,n)$ คือค่าความส่องสว่างของเฟรมปัจจุบันที่จุด (m,n)

$X^R(m,n)$ คือค่าความส่องสว่างของเฟรมอ้างอิงที่จุด (m,n)

2.4.2 ค่าอัตราส่วนสัญญาณยอดต่อสัญญาณรบกวน (Peak Signal to Noise Ratio: PSNR)

เป็นค่าที่วัดคุณภาพในแบบอัตนัย Objective ที่พยายามให้ความหมายในเชิงปรนัย Subjective [28] สามารถหาได้จากสมการที่ (2.31)

$$PSNR = 10 \log \left[\frac{R^2}{\frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N (X(m,n) - X^R(m,n))^2} \right] \quad (2.31)$$

โดยที่ $R = 2^L$ เมื่อ L แทนจำนวนบิต และ R แทนค่ายอดสูงสุดที่เป็นไปได้ของสัญญาณ

$M \times N$ คือขนาดของเฟรม

$X(m,n)$ คือค่าความส่องสว่างของเฟรมปัจจุบันที่จุด (m,n)

$X^R(m,n)$ คือค่าความส่องสว่างของเฟรมอ้างอิงที่จุด (m,n)

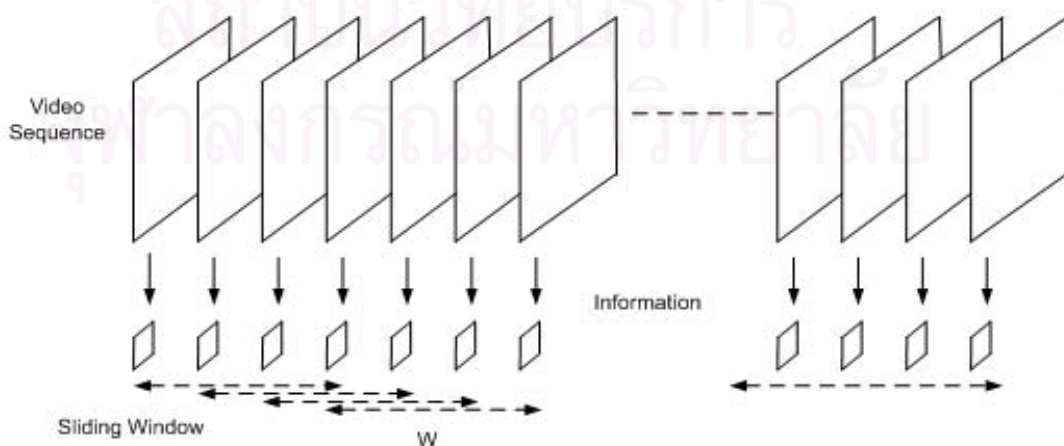
สำหรับวิทยานิพนธ์ฉบับนี้ใช้ค่าอัตราส่วนยอดต่อสัญญาณรบกวน (PSNR) เป็นค่าที่ใช้วัดประสิทธิภาพโดยเรียกว่าค่าคุณภาพสัญญาณวีดิทัศน์

2.5 การทบทวนงานวิจัยที่เกี่ยวข้อง

ในหัวข้อนี้จะกล่าวถึงงานวิจัยที่เกี่ยวข้องซึ่งเราจะประยุกต์นำมาใช้กับระเบียบวิธีการควบคุมอัตราที่เราจะนำเสนอ ได้แก่ ระเบียบวิธีการเข้ารหัสแบบหน้าต่างเลื่อน [29][33] และวิธีการของการประมาณบิตเป้าหมายแบบปรับตัวได้ [30]

2.5.1 ระเบียบวิธีการเข้ารหัสแบบหน้าต่างเลื่อน (A Sliding-Window Encoding Scheme) [29][33]

การเข้ารหัสสัญญาณวิดีโอที่อยู่นิ่งในแง่ของเวลาจริง เราสามารถที่จะนำหลักการของระเบียบวิธีการเข้ารหัสหน้าต่างเลื่อนนำมาใช้ในการควบคุมอัตราได้ โดยระเบียบวิธีการเข้ารหัสหน้าต่างเลื่อนจะใช้ข้อมูลทางสถิติของเฟรมในอนาคตมาช่วยในการจัดสรรบิต เมื่อ W เป็นขนาดของหน้าต่างเลื่อน เมื่อจะทำการเข้ารหัสเฟรมปัจจุบัน (เฟรม i) เราจะทำการเก็บข้อมูลทางสถิติของเฟรม $i+1, \dots, i+W-1$ ไว้ ข้อมูลเหล่านี้จะถูกใช้นำมาจัดสรรบิตใหม่เพื่อให้คุณภาพของสัญญาณวิดีโอดีขึ้น แสดงได้ดังรูปที่ ซึ่งข้อมูลทางสถิติเหล่านี้สามารถที่จะบอกได้ถึงความซับซ้อนของลำดับวิดีโอได้ การควบคุมอัตราก็จะจัดสรรบิตให้เหมาะสมตามความซับซ้อนของเฟรมนั้นๆ ซึ่งข้อมูลทางสถิติที่เก็บมานี้จะแตกต่างกันไป ซึ่งใน [29] จะใช้ค่าความแปรปรวนของเฟรมและจำนวนบิตข้อมูลส่วนหัว ใน [33] จะนำแบบจำลองอัตราและความผิดเพี้ยนโดเมน ρ (ρ -domain rate-distortion model) มาใช้ในการจัดสรรบิตที่เหมาะสมในระดับเฟรม ซึ่งค่าทางสถิติที่เก็บมาก็จะเป็นค่าพารามิเตอร์ต่างๆ ที่ใช้ในโดเมน ρ สำหรับในงานวิจัยที่นำเสนอ ข้อมูลทางสถิติก็จะต่างออกไป เนื่องจากทำบนพื้นฐานของมาตรฐาน H.264 ซึ่งรายละเอียดสามารถดูได้จากบทที่ 3



รูปที่ 2.18 หลักการของระเบียบวิธีการเข้ารหัสแบบหน้าต่างเลื่อน

มีหมายกำหนดการให้ถูกถอดรหัส ตัวถอดรหัสจะนำจำนวนบิตของเฟรมนั้น B_i ออกจากบัพเฟอร์เพื่อทำการถอดรหัส สามารถแสดงให้อยู่ในรูปรีเคอร์ซีฟสำหรับที่ช่วงเวลา $t = i / F$ ได้ดังนี้

$$\begin{cases} p_0 = P - B_0 \\ p_i = p(i/F) = p_{i-1} + R/F - B_i \quad \text{for } i = 1, 2, 3, \dots \end{cases} \quad (2.32)$$

เมื่อ B_i คือจำนวนของบิตสำหรับเฟรม i p_i คือจำนวนของบิตในบัพเฟอร์ตัวถอดรหัสหลังจากที่บิต B_i ถูกนำออกไปจากบัพเฟอร์แล้ว และ P คือจำนวนบิตที่ได้บรรจุก่อนในบัพเฟอร์ก่อนที่จะมีการถอดรหัส

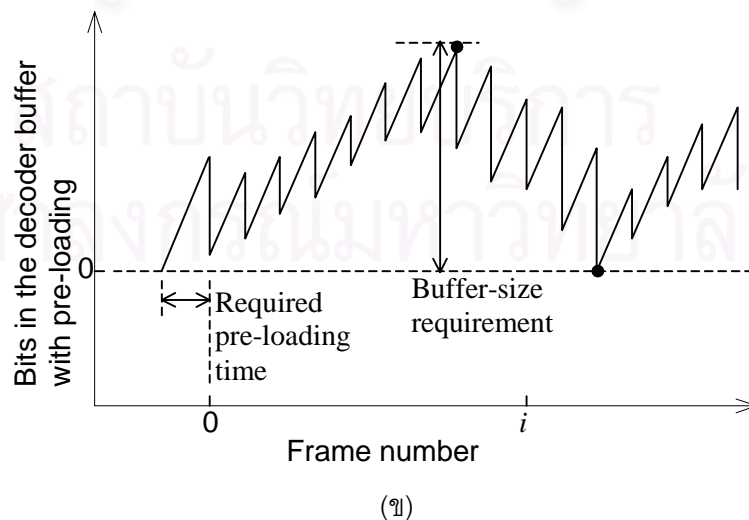
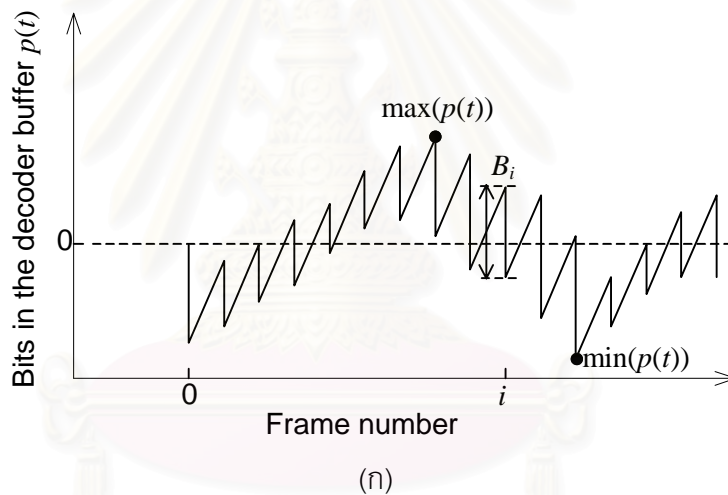
หลังจากมีการถอดรหัสเฟรม ถ้า $p_i > 0$ นั้นหมายความว่ายังมีบิตของเฟรมในอนาคตจำนวน p_i บิตที่ได้รับเข้ามาและถูกเก็บไว้เรียบร้อยแล้วในบัพเฟอร์ เพื่อที่จะป้องกันการล้นของข้อมูลในบัพเฟอร์ของตัวถอดรหัส ขนาดของบัพเฟอร์ (G) ควรจะมีขนาดใหญ่กว่าค่าสูงสุดของจำนวนบิตที่อยู่ในบัพเฟอร์ ($\max(p(t))$)

ถ้ามีเหตุการณ์ $p_i < 0$ เกิดขึ้นหลังจากที่มีการถอดรหัสเฟรม นั้นหมายความว่าไม่มีบิตเหลืออยู่ในบัพเฟอร์ (buffer underflow) และตัวถอดรหัสจะไม่มีบิตที่จำเป็นต้องใช้ในการถอดรหัสของเฟรมปัจจุบัน ในกรณีนี้ตัวถอดรหัสไม่สามารถที่จะถอดรหัสเฟรมที่เวลาหมายกำหนดการได้ และจำเป็นที่จะต้องรอสำหรับบิตที่เข้ามาซ้ำ การป้องกันไม่ให้มีสถานการณ์เช่นนี้เกิดขึ้น เราจำเป็นต้องทำการเพิ่มเวลาบรรจุก่อน L และเก็บบิตก้วหน้าเพิ่มขึ้น ตัวเข้ารหัสไม่ควรที่จะผลิตบิตสตรีมที่จะทำให้เกิด $p(t) < 0$

จากที่กล่าวมาเวลาบรรจุก่อน และขนาดบัพเฟอร์ตัวถอดรหัสเป็นพารามิเตอร์ที่สำคัญในงานประยุกต์วีดิทัศน์สตรีมมิงแบบออฟไลน์ในงานวิจัยที่ [29] และงานวิจัยที่นำเสนอจึงนำพารามิเตอร์เหล่านี้มาพิจารณาด้วย

2.5.2.1 การหาขนาดของบัฟเฟอร์ตัวถอดรหัสและเวลาบรรจุก่อนจากบิตสตรีมวีดิทัศน์ (Finding decoder buffer size and pre-loading time given a video bit-stream)

ขนาดของบัฟเฟอร์ตัวถอดรหัส และเวลาบรรจุก่อนที่เราต้องการจะสามารถคำนวณหาได้จากบิตสตรีมวีดิทัศน์ดังแสดงในรูปที่ 2.20 จากบิตสตรีมวีดิทัศน์จำนวนของบิต B_i ในแต่ละเฟรม i เราจะได้จากตัวเข้ารหัส กำหนดให้เสมือนว่าเรามีบัฟเฟอร์ของตัวถอดรหัสอยู่ และเฟรมแรก (เฟรม #0) ถูกถอดรหัสที่เวลา $t = 0$ จะสามารถเขียนกราฟได้ดังรูปที่ 2.20(ก) เมื่อ $p(t)$ คือบิตในบัฟเฟอร์ตัวถอดรหัสที่เวลา t เราจะสามารถหาค่าต่ำสุดและค่าสูงสุดของ $p(t)$ ได้จากกราฟเพื่อหลีกเลี่ยงไม่ให้เกิดการไม่มีข้อมูลในบัฟเฟอร์ $p(t)$ ควรจะมากกว่าหรือเท่ากับ 0 สำหรับทั้งลำดับวีดิทัศน์ ดังนั้นเราจะสามารถเลื่อนกราฟ $p(t)$ ขึ้นไปข้างบนด้วยค่าต่ำสุดของกราฟ ($\min(p(t))$) ค่าเวลาบรรจุก่อน และขนาดของบัฟเฟอร์จะสามารถหาได้จากกราฟดังรูปที่ 2.20(ข)



รูปที่ 2.20 การหาขนาดของบัฟเฟอร์ตัวถอดรหัส และเวลาบรรจุก่อนจากบิตสตรีมวีดิทัศน์

งานวิจัยที่นำเสนอซึ่งได้มีการพิจารณาเวลาบรรจุก่อน และขนาดของบัฟเฟอร์ตัวถอดรหัส ค่าพารามิเตอร์เหล่านี้สามารถหาได้จากบิตสตรีมวีดิทัศน์ได้เช่นเดียวกับหลักการในหัวข้อนี้

2.5.2.2 การผลิตบิตสตรีมวีดิทัศน์ที่ได้เข้ารหัสแล้วเมื่อมีการกำหนดขนาดของบัฟเฟอร์ตัวถอดรหัสและเวลาบรรจุก่อน (Generating a coded video bit-stream given the decoder buffer-size and pre-loading time)

ถ้ามีการกำหนดค่ามากที่สุดของเวลาบรรจุก่อน L และขนาดของบัฟเฟอร์ตัวถอดรหัส G ตัวเข้ารหัสควรจะผลิตบิตสตรีมที่ทำให้กราฟของบิตในบัฟเฟอร์ตัวถอดรหัส $p(t)$ ให้อยู่ในช่วง $[0, G]$ ในส่วนนี้เราจะพิจารณาว่าตัวเข้ารหัสต้องทำอะไรเพื่อที่จะปฏิบัติตามข้อกำหนดเหล่านี้

การป้องกันไม่ให้เกิดการไม่มีข้อมูลในบัฟเฟอร์ตัวถอดรหัส (ยกตัวอย่างเช่น ในสถานการณ์เมื่อตัวถอดรหัสมีกำหนดการตั้งถอดรหัสเฟรม i แต่บิตของเฟรม i นั้นยังไม่มีในบัฟเฟอร์) เราจำเป็นต้องทำให้ $p(t) \geq 0$ เสมอ จากสมการที่ (2.32) เราจะได้

$$P + \frac{R}{F} \cdot i - \sum_{k=0}^i B_k \geq 0 \quad (2.33)$$

เมื่อ P คือบิตที่บรรจุก่อนในบัฟเฟอร์ตัวถอดรหัส โดย $P = L \times R$

การป้องกันไม่ให้เกิดการล้นของข้อมูลในบัฟเฟอร์ตัวถอดรหัส (ยกตัวอย่างเช่น บิตที่เก็บไว้ในบัฟเฟอร์ตัวถอดรหัสมีค่าเกินกว่าขนาดของบัฟเฟอร์ตัวถอดรหัส) เราจำเป็นต้องทำให้ $p(t) \leq G$ เสมอ จากสมการที่ (2.32) เราจะได้

$$P + \frac{R}{F} \cdot (i+1) - \sum_{k=0}^i B_k \leq G \quad (2.34)$$

ถ้าตัวถอดรหัสมีขนาดบัฟเฟอร์ G บิต และตั้งเวลาบรรจุก่อน L วินาที (ยกตัวอย่างเช่น จำนวนบิตที่ถูกบรรจุก่อน $P = L \times R$ บิต) และตัวเข้ารหัสสามารถผลิตบิตสตรีมวีดิทัศน์ที่มีจำนวนบิต B_i สำหรับแต่ละเฟรม i (เฟรมแรกคือเฟรม 0) เป็นไปตามสมการที่ (2.33) และ (2.34) ดังนั้นบัฟเฟอร์ที่ตัวถอดรหัสจะไม่เกิดการไม่มีข้อมูลและการล้นของข้อมูลในบัฟเฟอร์

สำหรับในงานวิจัยที่ [29] เมื่อมีการกำหนดเวลาบรรจุก่อน และขนาดของบัฟเฟอร์ตัว ถอดรหัสมา ระเบียบวิธีการควบคุมอัตราสามารถที่จะควบคุมให้เป็นไปตามข้อกำหนดนั้นได้โดย การควบคุมบิตเป้าหมายไม่ให้มีจำนวนบิตในแต่ละเฟรมมีค่าที่จะทำให้เกิดการล้นของข้อมูลใน บัฟเฟอร์ตัวถอดรหัส และการไม่มีข้อมูลในบัฟเฟอร์ตัวถอดรหัส แต่สำหรับระเบียบวิธีที่นำเสนอเรา ไม่สามารถควบคุมอัตราให้เป็นไปตามข้อกำหนดจากบิตเป้าหมายได้ ซึ่งการแก้ปัญหาและวิธีการ ควบคุมอัตราให้เป็นไปตามข้อกำหนดนั้นรายละเอียดที่นำเสนอจะดูได้จากบทที่ 3

2.5.3 การประมาณบิตเป้าหมายแบบปรับตัวได้ [30]

สำหรับงานวิจัยในหัวข้อนี้จะเป็นระเบียบวิธีการควบคุมอัตราแบบส่ง ณ เวลาจริงบน มาตรฐานการเข้ารหัสวีดิทัศน์ H.264 โดยในส่วนควบคุมอัตราจะประมาณบิตเป้าหมายที่เหลืออยู่ ให้ปรับตัวได้ตามความซับซ้อนของเฟรมนั้นๆ โดยใช้อัตราส่วนค่าความแตกต่างสัมบูรณ์เฉลี่ย (MAD ratio) ซึ่งเป็นอัตราส่วนระหว่างค่าความแตกต่างสัมบูรณ์เฉลี่ยของเฟรมปัจจุบัน กับ ค่าเฉลี่ยของค่าความแตกต่างสัมบูรณ์เฉลี่ยของเฟรมตั้งแต่เริ่มต้นจนถึงเฟรมก่อนหน้าในกลุ่มของ ภาพ ดังสมการที่ (2.35)

$$MAD\ ratio = \frac{MAD_i}{MAD_{av}} \quad (2.35)$$

โดย $MAD_{av} = \frac{\sum_{k=0}^i MAD_k}{N_{pp}}$ คือค่าความแตกต่างสัมบูรณ์เฉลี่ยของเฟรมPทั้งหมดที่ได้รับการ เข้ารหัสแล้ว และ N_{pp} คือจำนวนเฟรมทั้งหมดที่ได้รับการเข้ารหัสแล้ว

ตัวเข้ารหัสจะทำการจัดสรรบิตสำหรับในแต่ละเฟรมตามค่าถ่วงน้ำหนักของอัตราส่วนค่า ความแตกต่างสัมบูรณ์เฉลี่ย ซึ่งเราสามารถนำมาประยุกต์ใช้กับระเบียบวิธีการควบคุมอัตราที่เรา นำเสนอได้ โดยจะประยุกต์นำมาใช้ในการเข้ารหัสครั้งที่สอง โดยมีการปรับเปลี่ยนค่าเพื่อให้ เหมาะสมกับระเบียบวิธีที่นำเสนอ รายละเอียดสามารถดูได้จากบทที่ 3

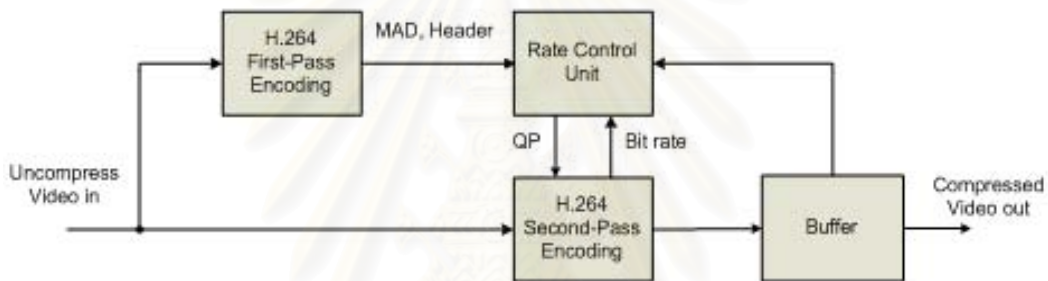
บทที่ 3

รายละเอียดเกี่ยวกับวิทยานิพนธ์และระบบที่นำเสนอ

วิทยานิพนธ์นี้นำเสนอเทคนิคการควบคุมอัตราสำหรับการเข้ารหัสวิดีโอที่สนับสนุนการบีบอัดแบบออนไลน์บนพื้นฐานของมาตรฐาน H.264 เพื่อให้คุณภาพของสัญญาณวิดีโอที่ส่งขึ้น รายละเอียดประกอบไปด้วยส่วนต่าง ๆ ดังต่อไปนี้

3.1 ภาพรวมของระบบ

ระบบจะประกอบไปด้วยส่วนต่างๆ ดังรูป 3.1



รูปที่ 3.1 ภาพรวมระบบ

ส่วนแรก คือลำดับภาพวิดีโอที่ไม่มีใครบีบอัดจะถูกเข้ารหัสครั้งแรกโดยมาตรฐานการเข้ารหัสสัญญาณวิดีโอ H.264 เพื่อเก็บข้อมูล คือค่าความแตกต่างสัมบูรณ์เฉลี่ย (MAD) ของเฟรม และจำนวนบิตของข้อมูลส่วนหัว (Header) ของเฟรม ข้อมูลเหล่านี้จะถูกนำไปใช้ช่วยในการเข้ารหัสครั้งที่สองในส่วนของการควบคุมอัตรา (Rate Control Unit) โดยเมื่อตัวเข้ารหัสกำหนดจำนวนบิตที่ต้องการให้กับหน่วยการควบคุมอัตรา หน่วยการควบคุมอัตราจะทำการหาค่าควอนไทซ์พารามิเตอร์ที่เหมาะสมเพื่อทำให้จำนวนบิตที่ต้องการสอดคล้องกับจำนวนบิตที่ออกสู่ช่องสัญญาณได้ และจำนวนบิตสะสมในบัฟเฟอร์ โดยจำนวนบิตที่ไม่สามารถออกสู่ช่องสัญญาณได้จะถูกเก็บไว้ในบัฟเฟอร์โดย ถ้ามีจำนวนบิตอยู่ในบัฟเฟอร์มาก หน่วยการควบคุมอัตราจะลดจำนวนบิตที่เกิดจากตัวลงรหัสเพื่อป้องกันบัฟเฟอร์เต็ม ในกรณีถ้ามีจำนวนบิตอยู่ในบัฟเฟอร์น้อย หน่วยการควบคุมอัตราจะเพิ่มจำนวนบิตที่เกิดจากตัวลงรหัสเพื่อป้องกันกรณีไม่มีข้อมูลในบัฟเฟอร์

3.2 ลักษณะสัญญาณวิดีโอที่สนทนาเข้า

สัญญาณวิดีโอที่สนทนาเข้าที่ใช้ในวิทยานิพนธ์นี้ คือลำดับภาพมาตรฐาน (Standard sequence) ที่อยู่ในรูปแบบ QCIF (176x144 จุดภาพ) ซึ่งในแต่ละลำดับภาพจะมีการเคลื่อนไหวที่มาก และเร็ว โดยลำดับภาพ Foreman มีการเลื่อนและหมุนกลิ้ง ตัวคนมีการหมุนคอ เอียงคอ ลำดับภาพ Carphone เป็นการบันทึกในรถขณะเคลื่อนที่ ซึ่งมีลักษณะคล้ายฉากในภาพยนตร์ที่เหมาะสมกับการทดสอบแบบสตรีมมิง โดยลำดับภาพ Suzie+Trevor คือลำดับภาพมาตรฐานสองชนิดที่นำมาต่อกัน (Suzie และ Trevor) เพื่อทดสอบสถานการณ์การเปลี่ยนของฉาก ดังรูปที่ 3.2



(ก)



(ข)



(ค)



(ง)

รูปที่ 3.2 สัญญาณวิดีโอที่สนทนาที่ใช้ในวิทยานิพนธ์

(ก) Foreman มีจำนวนเฟรมเท่ากับ 400

(ข) Carphone มีจำนวนเฟรมเท่ากับ 381

(ค) Suzie มีจำนวนเฟรมเท่ากับ 150

(ง) Trevor มีจำนวนเฟรมเท่ากับ 150

3.3 การทำงานของระบบที่นำเสนอ

ในงานประยุกต์วีดิทัศน์สตรีมมิ่งสามารถแบ่งออกเป็นสองประเภท คือแบบออนไลน์ หรือแบบส่ง ณ เวลาจริง และแบบออฟไลน์ หรือแบบไม่ส่ง ณ เวลาจริง โดยในแบบออนไลน์ ลำดับวีดิทัศน์จะถูกเข้ารหัสเพียงรอบเดียว (one-pass encoding) เพื่อให้ทันส่งในเวลาจริง การประวิงเวลาจึงเป็นประเด็นสำคัญสำหรับการสื่อสารในเวลาจริง แต่สำหรับงานประยุกต์วีดิทัศน์สตรีมมิ่งแบบออฟไลน์ ลำดับวีดิทัศน์จะถูกเข้ารหัสแบบออฟไลน์เพื่อเก็บ (stored) ไว้ จึงไม่ต้องพิจารณาเวลาประวิงของตัวเข้ารหัส เวลาบรรจุก่อนในบัฟเฟอร์ตัวถอดรหัส (pre-loaded time) ก่อนที่จะมีการแสดงที่ปลายทาง และขนาดของบัฟเฟอร์ตัวถอดรหัสเป็นสิ่งที่ต้องนำมาพิจารณา สำหรับการดำเนินงานของระบบที่นำเสนอนี้เนื่องจากเป็นงานประยุกต์วีดิทัศน์แบบออฟไลน์ ลำดับวีดิทัศน์จะถูกเข้ารหัสสองรอบ (two-pass encoding) โดยในรอบแรกของการเข้ารหัสจะทำเพื่อเก็บข้อมูล คือค่าความแตกต่างสัมบูรณ์เฉลี่ย (MAD) ของเฟรม และจำนวนบิตของข้อมูลส่วนหัว (header) ของเฟรม ข้อมูลเหล่านี้จะถูกนำไปใช้ช่วยในการเข้ารหัสรอบที่สองในส่วนของการควบคุมอัตรา โดยในการเข้ารหัสรอบที่สองจะประยุกต์ใช้ ระเบียบวิธีการเข้ารหัสที่พิจารณาข้อมูลทางสถิติของเฟรมในอนาคตด้วยหน้าต่างเลื่อน (sliding window) [29][33] และวิธีการประมาณบิตเป้าหมายแบบปรับตัวได้ตามความซับซ้อนของการเข้ารหัสในแต่ละเฟรม โดยการใช้อัตราส่วนค่าความแตกต่างสัมบูรณ์เฉลี่ย (MAD ratio) [30] การเปลี่ยนของฉาก และเฟรมที่มีการเคลื่อนไหวที่มาก ๆ จะถูกนำมาพิจารณาเพื่อให้มีคุณภาพที่ดีขึ้น คุณภาพของวีดิทัศน์โดยรวมจะดีขึ้นเพราะว่าตัวเข้ารหัสจะใช้ประโยชน์จากสถิติของเฟรมในอนาคต โดยจะจัดสรรบิตมากกับเฟรมที่มีความซับซ้อนมากและบิตน้อยกับเฟรมที่มีความซับซ้อนน้อย

3.3.1 ระเบียบวิธีการควบคุมอัตราที่นำเสนอ

สำหรับระเบียบวิธีการควบคุมอัตราที่นำเสนอนี้จะประยุกต์ใช้ระเบียบวิธีการเข้ารหัสแบบหน้าต่างเลื่อน และการประมาณบิตเป้าหมายแบบปรับตัวได้ โดยจะอยู่บนพื้นฐานของระเบียบวิธีการควบคุมอัตราของมาตรฐาน H.264 ซึ่งขั้นตอนการประมาณบิตเป้าหมายของเฟรมในระเบียบวิธีการควบคุมอัตราในมาตรฐาน H.264 จากสมการที่ (2.11) บิตเป้าหมายที่เหลืออยู่ (T_r) จะถูกจัดสรรให้กับเฟรมทั้งหมดที่ยังไม่ได้รับการเข้ารหัสเท่า ๆ กัน ระเบียบวิธีการควบคุมอัตราที่เราจะนำเสนอจึงมองเฉพาะเจาะจงลงไปทีบิตเป้าหมายที่เหลืออยู่ นั่นคือในขั้นตอนของการประมาณบิตเป้าหมายเริ่มต้น บิตที่เหลืออยู่ควรจะถูกจัดสรรให้กับเฟรมทั้งหมดที่ยังไม่ได้รับการเข้ารหัสตามความซับซ้อนของเฟรมนั้นๆ โดยระเบียบวิธีการควบคุมอัตราที่ใช้ในมาตรฐาน

H.264 ในขั้นตอนการหาค่าควอนไทซ์พารามิเตอร์ จะหาได้โดยใช้โมเดลกำลังสอง (quadratic R-Q model) [16] ซึ่งโมเดลกำลังสองนี้ ความซับซ้อนในการเข้ารหัสจะถูกแทนด้วยค่าความแตกต่างสัมบูรณ์เฉลี่ย (MAD) ซึ่งค่าความแตกต่างสัมบูรณ์เฉลี่ยนี้เป็นตัวบ่งชี้ที่ดีของความซับซ้อนของเฟรม [26] ดังนั้นเราจึงใช้ค่าความแตกต่างสัมบูรณ์เฉลี่ยของเฟรมนี้มาใช้ในการจัดสรรบิตตามความซับซ้อนของเฟรมนั้น ๆ

3.3.1.1 การเข้ารหัสครั้งแรก (First-Pass Encoding)

การเข้ารหัสครั้งแรกทำเพื่อรวบรวมข้อมูล ลำดับภาพวิดีโอที่ต้นจะถูกเข้ารหัสโดยมาตรฐาน H.264 โดยใช้ค่าควอนไทซ์ที่คงที่ ถ้าตัวเข้ารหัสทำการเข้ารหัสเฟรม i ข้อมูลต่าง ๆ คือค่าความแตกต่างสัมบูรณ์เฉลี่ย (MAD) ของเฟรม และจำนวนบิตสำหรับข้อมูลส่วนหัว (header) ของเฟรม $i+1, \dots, i+W-1$ จะถูกเก็บไว้ ซึ่งประกอบไปด้วยบิตของเวกเตอร์การเคลื่อนที่ บิตของชนิดของโหมดที่เลือกในการเข้ารหัส ข้อมูลเหล่านี้จะถูกใช้ในการเข้ารหัสในครั้งที่สองเพื่อให้คุณภาพของสัญญาณวิดีโอที่ดีขึ้น

3.3.1.2 การเข้ารหัสครั้งที่สอง (Second-Pass Encoding)

การเข้ารหัสครั้งที่สองจะประยุกต์ใช้ระเบียบวิธีการเข้ารหัสแบบหน้าต่างเคลื่อน และการประมาณบิตเป้าหมายแบบปรับตัวได้ โดยบิตจะถูกจัดสรรให้กับเฟรมตามความซับซ้อนของเฟรมนั้น ๆ ตัวเข้ารหัสจะทำการจัดสรรบิตสำหรับในแต่ละเฟรมตามค่าถ่วงน้ำหนักของค่าความแตกต่างสัมบูรณ์เฉลี่ย (MAD_i) ของเฟรม i ในจำนวน W เฟรม ตามขั้นตอนดังนี้

ขั้นตอนที่ 1: การกำหนดค่าเริ่มต้น ของตัวคูณค่าจำนวนบิต $\delta_{-1} = 0$ เมื่อ δ_i ใช้ในการสังเกตการณ์ความแตกต่าง (เกินหรือขาด) ระหว่างจำนวนบิตเป้าหมายอุดมคติและจำนวนบิตแท้จริงที่ถูกสร้างมาจากตัวเข้ารหัสที่เฟรม i

ขั้นตอนที่ 2: การคำนวณบิตเป้าหมายที่เหลืออยู่ T_r สำหรับเฟรม i สามารถคำนวณหาได้จาก ($i = 0, 1, 2, \dots$)

$$T_r = \left(\frac{R}{F} \times W - \sum_{k=i}^{i+W-1} H_k - \delta_{i-1} \right) \times \frac{MAD_i}{\sum_{k=i}^{i+W-1} MAD_k} + H_i \quad (3.1)$$

เมื่อ R คืออัตราส่วนของสัญญาณ F คืออัตราเฟรม W คือขนาดหน้าต่าง และ H_i คือจำนวนของบิตสำหรับข้อมูลส่วนหัวของเฟรม i

ขั้นตอนที่ 3: การประมาณบิตเป้าหมายที่เหลืออยู่แบบปรับตัวได้ เราจะประยุกต์ใช้อัตราส่วนค่าความแตกต่างสัมบูรณ์เฉลี่ย (MAD ratio) ในการควบคุมบิตเป้าหมายที่เหลืออยู่อีกครั้ง ซึ่งเป็นอัตราส่วนระหว่างค่าความแตกต่างสัมบูรณ์เฉลี่ยของเฟรมปัจจุบัน กับค่าเฉลี่ยของค่าความแตกต่างสัมบูรณ์เฉลี่ยของเฟรมตั้งแต่เริ่มต้นจนถึงเฟรมก่อนหน้าในกลุ่มของภาพ ในการคำนวณเราจะใช้กฎดังต่อไปนี้ในการปรับบิตเป้าหมายที่เหลืออยู่

คำนวณค่าเฉลี่ยของค่าความแตกต่างสัมบูรณ์เฉลี่ย (Average MAD) ของเฟรมทั้งหมดที่ได้รับการเข้ารหัสแล้วตามสมการที่ (3.2)

$$MAD_{av} = \frac{\sum_{k=0}^i MAD_k}{N_{pp}} \quad (3.2)$$

เมื่อ N_{pp} คือจำนวนเฟรมทั้งหมดที่ได้รับการเข้ารหัสแล้ว

คำนวณอัตราส่วนค่าความแตกต่างสัมบูรณ์เฉลี่ย (MAD ratio) คือค่าความแตกต่างสัมบูรณ์เฉลี่ยของเฟรมปัจจุบันหารด้วยค่าเฉลี่ยของค่าความแตกต่างสัมบูรณ์เฉลี่ยของเฟรมตั้งแต่เริ่มต้นจนถึงเฟรมก่อนหน้าในกลุ่มของภาพ

$$MAD \text{ ratio} = \frac{MAD_i}{MAD_{av}} \quad (3.3)$$

ถ้า ($MAD \text{ ratio} < 0.7$) จะได้

$$T_r = T_r * 0.5 \quad (3.4)$$

หากนอกจากนั้น ถ้า ($MAD \text{ ratio} < 1.0$) จะได้

$$T_r = T_r * MAD \text{ ratio} * 0.6 \quad (3.5)$$

หากนอกจากนั้น ถ้า ($MAD \text{ ratio} < 1.9$) จะได้

$$T_r = T_r * MAD \text{ ratio} * 0.7 \quad (3.6)$$

หากนอกจากนั้น ถ้า ($MAD\ ratio \geq 1.9$) จะได้

$$T_r = T_r * 1.5 \quad (3.7)$$

ในระหว่างนั้น เราจะทำการเพิ่มค่า β เป็น 0.7 จาก 0.5 ในสมการที่ (2.11) ดังนั้น บิตเป้าหมายที่เหลืออยู่ (T_r) จะมีการถ่วงน้ำหนักมากกว่าบิตเป้าหมายของบัฟเฟอร์ (T_{buf}) ค่าพารามิเตอร์ต่างๆ ที่ถูกใช้ในกฎข้างบนมาจากที่ตั้งขึ้นไว้โดยอาศัยความสังเกตจากการทดลอง แนวความคิดพื้นฐานคือการให้บิตเป้าหมายที่เหลืออยู่มีค่าน้อยกว่าถ้าเฟรมปัจจุบันมีความซับซ้อนน้อย และการให้บิตเป้าหมายที่เหลืออยู่มีค่ามากกว่าถ้าเฟรมปัจจุบันมีความซับซ้อนมาก จุดมุ่งหมายของการปรับปรุงนี้คือการประหยัดบิตจากเฟรมที่มีความซับซ้อนน้อย และจัดสรรบิตให้เพิ่มมากขึ้นกับเฟรมที่มีความซับซ้อนมากเนื่องมาจากในเฟรมมีการเคลื่อนไหวสูงหรือมีการเปลี่ยนแปลงของฉาก ในที่สุดค่าบิตเป้าหมายสุดท้าย (T) จะถูกคำนวณตามสมการที่ (2.11)

ขั้นตอนที่ 4: การปรับตัวคุมค่าจำนวนบิต หลังจากทำการเข้ารหัสในแต่ละเฟรม บิตที่แท้จริงที่ใช้ในการเข้ารหัสเฟรม i จะถูกนำมาปรับตัวคุมค่าจำนวนบิต ตัวคุมค่าจำนวนบิตจะถูกปรับโดย $\delta_i = \delta_{i-1} + T' - R/F$ ถ้ามีเฟรมที่จะเข้ารหัสเพิ่มจะกลับไปทำขั้นตอนที่สอง หรือถ้าไม่มีเฟรมที่จะเข้ารหัสแล้วก็จะหยุด

3.3.1.3 ระเบียบวิธีการควบคุมอัตราที่นำเสนอกับข้อกำหนดของขนาดบัฟเฟอร์ตัวถอดรหัสและเวลาบรรจุก่อน

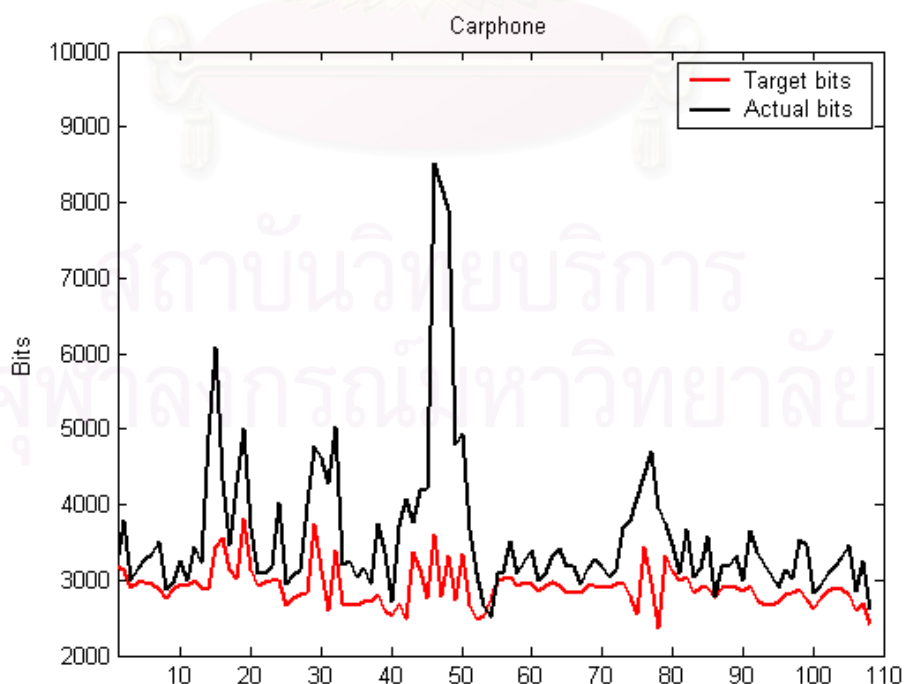
สำหรับงานประยุกต์วีดิทัศน์สตรีมมิงแบบออฟไลน์ ลำดับวีดิทัศน์จะถูกเข้ารหัสแบบออฟไลน์เพื่อเก็บไว้ จึงไม่ต้องพิจารณาเวลาประวิงของตัวเข้ารหัส เวลาบรรจุก่อนในบัฟเฟอร์ตัวถอดรหัสก่อนที่จะมีการแสดงที่ปลายทาง และขนาดของบัฟเฟอร์ตัวถอดรหัสเป็นสิ่งที่ต้องนำมาพิจารณา โดยขนาดของบัฟเฟอร์ตัวถอดรหัส และเวลาบรรจุก่อนที่ต้องการของบิตสตรีมวีดิทัศน์ที่เข้ารหัสโดยใช้ระเบียบวิธีการควบคุมอัตราที่นำเสนอสามารถจะหาได้เช่นเดียวกับในหัวข้อที่ 2.5.2.1

ในงานประยุกต์วีดิทัศน์สตรีมมิงแบบออฟไลน์ วีดิทัศน์ที่เข้ารหัสแล้วที่ถูกเก็บไว้ในเครื่องแม่ข่ายจะถูกส่งไปยังตัวถอดรหัสผ่านทางช่องสัญญาณสื่อสาร บัฟเฟอร์ที่ตัวถอดรหัสจะมีความจำเป็นในการเก็บบิตที่เกินมาที่ได้รับมาจากช่องสัญญาณเพื่อรอถอดรหัสให้แสดงตรงตามเวลาที่

ปลายทาง บิตของเฟรมในอนาคตสามารถส่งก้าวหน้าไปยังบัฟเฟอร์ตัวถอดรหัสได้เรื่อยๆ ตรงเท่าที่ช่อง สัญญาณมีแบนวิดท์ และไม่มีการล้นของข้อมูลในบัฟเฟอร์ที่ตัวถอดรหัส (buffer overflow) โดยที่ผู้ใช้จะเป็นผู้กำหนดว่าเวลาการบรรจุก่อนควรมีค่าเท่าใดขึ้นอยู่กับการถอดทวนของผู้ใช้ ซึ่งถ้ามีการกำหนดขนาดของบัฟเฟอร์ที่ตัวถอดรหัส และค่าเวลาบรรจุก่อนจากผู้ใช้ ระเบียบวิธีการควบคุมอัตราที่จะนำเสนอต่อไปจะสามารถผลิตบิตสตรีมวิดิทัศน์ซึ่งสามารถเป็นไปตามข้อกำหนดได้

กำหนดให้ตัวเข้ารหัสต้องการผลิตบิตสตรีมวิดิทัศน์สำหรับตัวถอดรหัสที่มีขนาดบัฟเฟอร์ G และผู้ใช้อนุญาตให้มีค่าประวิงเวลาบรรจุก่อน L วินาที สำหรับช่องสัญญาณอัตราคงที่ R ค่าบิตมากที่สุดที่สามารถบรรจุก่อนในบัฟเฟอร์ตัวถอดรหัสมีค่า $P = R \times L$ ในหัวข้อนี้เราจะอธิบายถึงระเบียบวิธีการควบคุมอัตราที่เป็นไปตามข้อกำหนดเหล่านี้

จากการศึกษาและการทดลองพบว่า บิตเป้าหมายจากระเบียบวิธีการควบคุมอัตราในมาตรฐาน H.264 จะมีค่าแตกต่างจากบิตที่ใช้ในการเข้ารหัสจริงมาก ดังภาพที่ 3.3 เนื่องมาจากค่าควอนไทซ์พารามิเตอร์ถูกจำกัดไม่ให้มีค่าเกินกว่า 3 ระดับในเฟรมที่ติดกัน ทำให้ไม่สามารถควบคุมอัตราให้เป็นไปตามข้อกำหนดของขนาดบัฟเฟอร์ตัวถอดรหัส และเวลาบรรจุก่อนโดยใช้ระเบียบวิธีตาม [29] ได้



รูปที่ 3.3 จำนวนบิตเป้าหมาย และจำนวนบิตที่ลงรหัสจริงสำหรับลำดับภาพ Carphone

จึงได้นำเสนอการกำหนดเงื่อนไขการเลือกค่าควอนไทล์พารามิเตอร์ โดยคำนึงถึงจำนวนบิตสะสมทั้งหมดในขณะนั้น (B_k) หลังจากลงรหัสหน่วยพื้นฐานย่อย k

กำหนดให้ B_k คือจำนวนบิตทั้งหมดของเฟรมที่ได้รับการเข้ารหัสแล้วรวมกับจำนวนบิตหลังจากเข้ารหัสหน่วยพื้นฐานย่อย k ที่เฟรม i ดังแสดงในสมการ

$$B_k = \sum_{j=0}^i B_j + \sum_{i=1}^k B_{b-i} \quad (3.8)$$

เมื่อ B_j คือจำนวนบิตของเฟรมที่ได้รับการเข้ารหัสแล้วตั้งแต่เฟรมแรกจนถึงเฟรมที่ i และ B_{b-i} คือจำนวนบิตที่เข้ารหัสแล้วตั้งแต่หน่วยพื้นฐานย่อยที่ 1 ถึง k ที่เฟรม i

ในขณะขั้นตอนการเข้ารหัสครั้งที่สอง ตามสมการที่ (3.1) และ (3.2) เพื่อหลีกเลี่ยงไม่ให้เกิดการไม่มีข้อมูล และการล้นของข้อมูลในบัฟเฟอร์ตามข้อกำหนดที่กำหนดไว้จำเป็นจะต้องเป็นไปตามเงื่อนไขในหัวข้อที่ 2.5.2.2 ตามสมการที่ (2.33) และ (2.34) จากค่า B_k เราจะเช็คได้ว่าบัฟเฟอร์ตัวถอดรหัสเริ่มที่จะเข้าใกล้สถานการณ์ที่ไม่มีข้อมูล และมีการล้นของข้อมูลในบัฟเฟอร์ (ด้วยสมการที่ (3.9) และ (3.10)) เมื่อ α_1 และ α_2 คือค่าขีดเริ่มคงที่ ในการทดลอง เรากำหนดให้ $\alpha_1 = 1.2$ และ $\alpha_2 = 0.8$ ค่าขีดเริ่มคงที่ α_1 และ α_2 จะถูกใช้เป็นขีดเฝ้าระวังว่าเริ่มที่จะเข้าใกล้สถานการณ์ที่ไม่มีข้อมูล และมีการล้นของข้อมูลในบัฟเฟอร์ตามลำดับ ตามสมการดังนี้

$$\text{ถ้า } \alpha_1 \cdot B_k \geq P + \frac{R}{F} \cdot i$$

เราจะทำการเพิ่มค่าควอนไทล์พารามิเตอร์ของหน่วยพื้นฐานย่อยตามสมการ

$$Q_{cb} = Q_{cb} + 2 \quad (3.9)$$

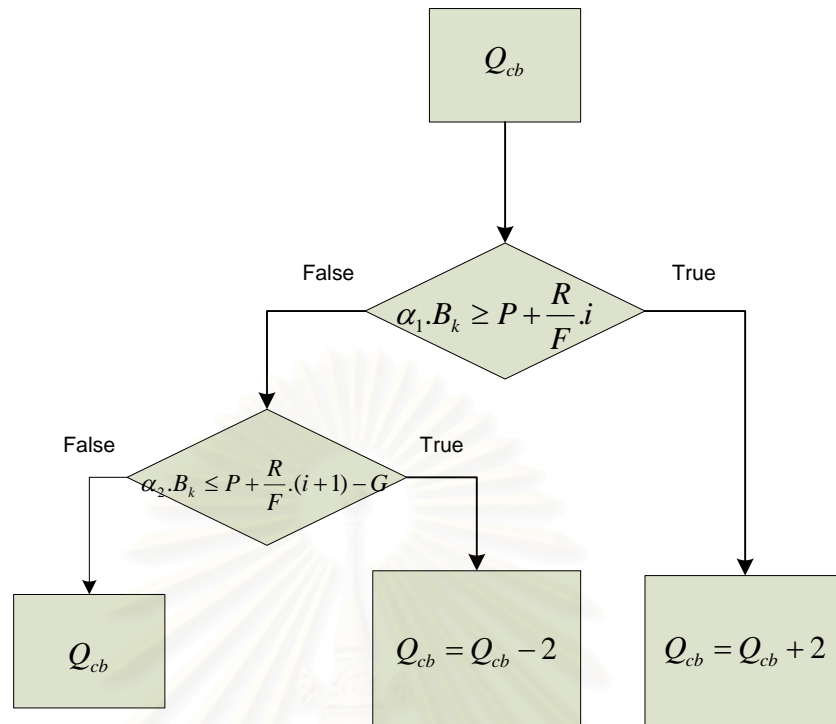
$$\text{หากนอกจากนั้น ถ้า } \alpha_2 \cdot B_k \leq P + \frac{R}{F} \cdot (i+1) - G$$

เราจะทำการลดค่าควอนไทล์พารามิเตอร์ของหน่วยพื้นฐานย่อยตามสมการ

$$Q_{cb} = Q_{cb} - 2 \quad (3.10)$$

หากนอกจากนั้น ค่าควอนไทล์ของหน่วยพื้นฐานย่อยจะมีค่าเท่าเดิม Q_{cb}

เมื่อ Q_{cb} คือค่าควอนไทล์พารามิเตอร์ของหน่วยพื้นฐานย่อยปัจจุบัน สามารถแสดงดังรูปที่ 3.4



รูปที่ 3.4 ระเบียบวิธีการควบคุมอัตราที่นำเสนอกับข้อกำหนดของขนาดบัฟเฟอร์ตัวถอดรหัสและเวลาบรรจุก่อน

บทที่ 4

ผลการทดสอบ

ในบทนี้จะเป็นการทดสอบระบบที่นำเสนอ โดยส่วนแรกจะกล่าวถึงรายละเอียดในการทดสอบ ผลการทดสอบและการวิเคราะห์ประสิทธิภาพของการเข้ารหัสวีดิทัศน์ที่นำเสนอ ซึ่งอยู่ในหัวข้อ 4.1 จากนั้น หัวข้อ 4.2 คือ ผลการทดสอบ และ ทำการวิเคราะห์และวิจารณ์ผลการทดสอบต่าง ๆ

4.1 ขั้นตอนการทดสอบ

สำหรับเทคนิคการควบคุมอัตราสำหรับการเข้ารหัสวีดิทัศน์สตรีมมิงแบบฮอปไลน์บนพื้นฐานของมาตรฐาน H.264 จะแบ่งขั้นตอนการทดสอบออกเป็นขั้นตอนต่าง ๆ ดังนี้

- เมื่อไม่มีข้อกำหนดของขนาดบัฟเฟอร์ตัวถอดรหัส และเวลาบรรจุก่อน จะทำการทดสอบระบบที่นำเสนอในแง่ของคุณภาพของสัญญาณวีดิทัศน์เปรียบเทียบกับมาตรฐาน H.264 ที่อัตราบิตเป้าหมายต่าง ๆ กัน คือที่ 32 64 128 และ 256 กิโลบิตต่อวินาที และที่ขนาดหน้าต่างที่ต่างกัน คือที่ $W = 50$ และ 100 และทำการหาค่าขนาดบัฟเฟอร์ตัวถอดรหัสที่ต้องการ และค่าเวลาบรรจุก่อนที่ต้องการ สำหรับแต่ละลำดับสัญญาณวีดิทัศน์
- เมื่อมีข้อกำหนดของขนาดบัฟเฟอร์ตัวถอดรหัส และเวลาบรรจุก่อน จะทำการทดสอบระบบที่นำเสนอในแง่ของคุณภาพของสัญญาณวีดิทัศน์เปรียบเทียบกับมาตรฐาน H.264 ที่อัตราบิตเป้าหมายต่าง ๆ กัน คือที่ 32 64 128 และ 256 กิโลบิตต่อวินาที และที่ขนาดหน้าต่างที่ต่างกัน คือที่ $W = 50$ และ 100 ของลำดับสัญญาณวีดิทัศน์ Suzie+Trevor ซึ่งเป็นลำดับภาพมาตรฐานสองชนิดที่นำมาต่อกัน (Suzie และ Trevor) เพื่อทดสอบสถานการณ์การเปลี่ยนของฉาก

4.1.1 การกำหนดพารามิเตอร์ในการเข้ารหัส

1. โปรแกรมที่ใช้ในการเข้ารหัสมาตรฐาน H.264 และระบบที่นำเสนอ คือ JM6.1 (Joint Test Model 6.1) [31] ซึ่งเป็นแบบจำลองทดสอบร่วม (Joint Test Model) สร้างขึ้นโดย Joint Video Team (JVT) เพื่อใช้ในการทดสอบการเข้ารหัสมาตรฐาน H.264

2. การกำหนดอัตราเฟรมในการเข้ารหัส ในการทดลองนี้กำหนดใช้อัตราเฟรมเท่ากับ 10 เฟรมต่อวินาที ซึ่งเป็นอัตราเฟรมที่ใช้ในการลงรหัสสัญญาณวิดีโอที่อัตราบิตต่ำ และจำนวนเฟรมที่เข้ารหัสเท่ากับ 100 เฟรม
3. การกำหนดชนิดของสัญญาณวิดีโอที่ที่ใช้ในการลงรหัส ซึ่งในที่นี้เป็น QCIF 4:2:0 มีจำนวนจุดภาพ 176x144 จุดภาพ
4. การเลือกเฟรมอ้างอิง (Intra Frame Selection) ใช้เฟรมแรกเพียงเฟรมเดียวที่เป็น I เฟรมตลอดการลงรหัสลำดับภาพเนื่องจากการลงรหัสเฟรมชนิดนี้ต้องใช้บิตจำนวนมาก
5. กำหนดอัตราบิตเป้าหมายในการส่งที่ 32 64 128 และ 256 กิโลบิตต่อวินาที
6. ค่าควอนไทซ์สำหรับ I เฟรมแรกเท่ากับ 28, 32 และ 36 สำหรับลำดับสัญญาณวิดีโอที่ศน์ Suzie+Trevor Carphone และ Foreman ตามลำดับ

ค่าพารามิเตอร์อื่น ๆ ที่ใช้ในการทดลองดังแสดงในตารางที่ 4.1

ตารางที่ 4.1 พารามิเตอร์ที่ใช้ในการทดสอบ

MV resolution	¼ pel
Hadamard	on
RD optimization	on
Search Range	±32(QCIF)
Reference frames	1
Restrict Search Range	2
Symbol mode	UVLC
GOP structure	IPPP

4.1.2 ลำดับวิดีโอที่ที่ใช้ในการทดสอบ

โดยลักษณะทั่ว ๆ ไปของลำดับวิดีโอที่ที่ใช้ในการทดสอบ มีดังนี้

1. ลำดับวิดีโอที่ศน์ Foreman มีการเคลื่อนที่ของวัตถุด้วยความเร็วปานกลาง มีการเลื่อนและหมุนกล้อง ตัวคนมีการหมุนคอ เอียงคอ

2. ลำดับวีดิทัศน์ Carphone มีการเคลื่อนที่ของวัตถุเร็วเพราะเป็นภาพที่บันทึกบนรถยนต์ขณะเคลื่อนที่ มีการเปลี่ยนแปลงของพื้นหลัง และเคลื่อนที่ของคน
3. ลำดับวีดิทัศน์ Suzie+Trevor ซึ่งเป็นลำดับภาพมาตรฐานสองชนิดที่นำมาต่อกัน (Suzie และ Trevor) เพื่อทดสอบสถานการณ์การเปลี่ยนของฉาก โดยลำดับวีดิทัศน์ Suzie มีการเคลื่อนที่บริเวณศีรษะและไหล่ มีการก้มและเงยหน้า ส่วนลำดับวีดิทัศน์ Trevor มีวัตถุภายในเฟรมเดียวกันหลายวัตถุซึ่งเคลื่อนที่ต่างกันออกไป และมีการตัดภาพ

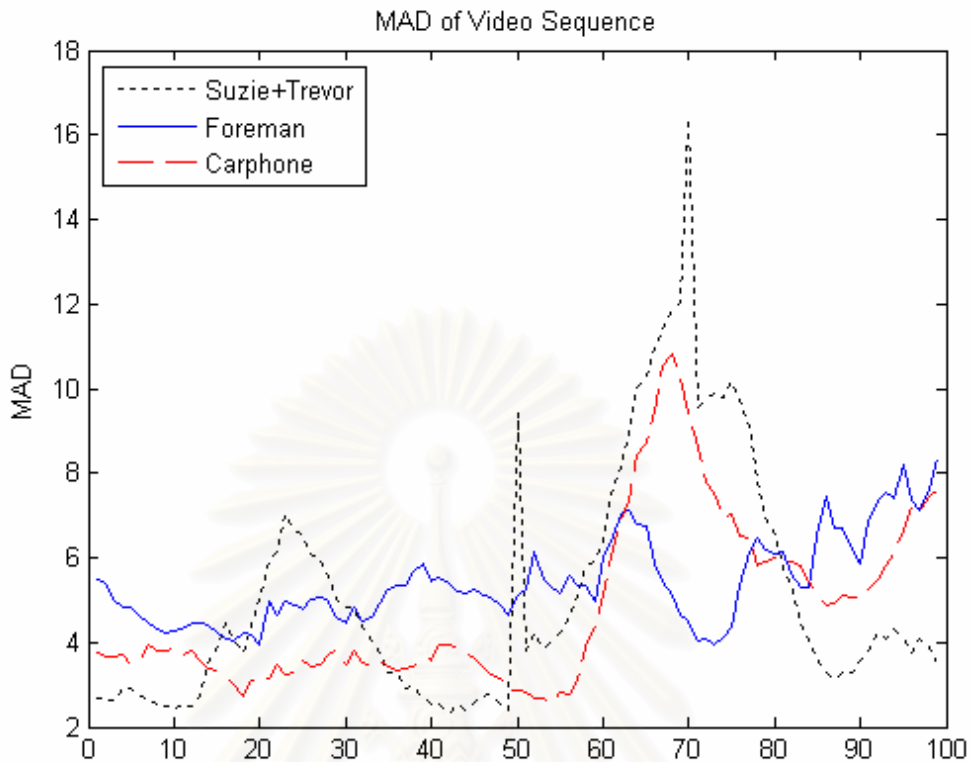
4.2 ผลการทดสอบและวิเคราะห์ผลการทดสอบ

สำหรับเทคนิคการควบคุมอัตราสำหรับการเข้ารหัสวีดิทัศน์สตรีมมิงแบบฮอปไลน์บนพื้นฐานของมาตรฐาน H.264 จะทำการทดสอบในเรื่องของคุณภาพของสัญญาณวีดิทัศน์เปรียบเทียบกับมาตรฐาน H.264 ดังขั้นตอนต่อไปนี้

4.2.1 การเข้ารหัสครั้งแรก

ลำดับภาพวีดิทัศน์จะถูกเข้ารหัสครั้งแรกโดยใช้ค่าควอนไทซ์ที่คงที่ ถ้าตัวเข้ารหัสทำการเข้ารหัสเฟรม i ข้อมูลต่างๆ คือค่าความแตกต่างสัมบูรณ์เฉลี่ย (Mean Absolute Difference) และจำนวนบิตสำหรับข้อมูลส่วนหัวของเฟรม $i + 1, \dots, i + W - 1$ จะถูกเก็บไว้ ซึ่งประกอบไปด้วยบิตของเวกเตอร์การเคลื่อนที่ บิตของชนิดของโหมดที่เลือกในการเข้ารหัส ฯลฯ ข้อมูลเหล่านี้สามารถนำมาใช้เป็นตัวช่วยในการจัดสรรบิตในการเข้ารหัสครั้งที่สอง และทำให้คุณภาพของวีดิทัศน์ดีขึ้น ดังรูปที่ 4.1 คือค่าความแตกต่างสัมบูรณ์เฉลี่ยของลำดับวีดิทัศน์ทั้งสามที่ได้มากจากขั้นตอนการเข้ารหัสครั้งแรกนี้

จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 4.1 ค่าความแตกต่างสัมบูรณ์เฉลี่ยของลำดับวิดีโอที่ใช้ในการทดสอบ

จากกราฟ สำหรับลำดับวิดีโอ Foreman ซึ่งมีการเคลื่อนไหวของวัตถุด้วยความเร็วปานกลาง ในแต่ละเฟรมจึงมีความซับซ้อนไม่แตกต่างกันมาก ค่าความแตกต่างสัมบูรณ์เฉลี่ยที่ได้จึงค่อนข้างเกาะกลุ่มกันทั้งลำดับวิดีโอ แต่ในช่วงเฟรมที่ 60 จนถึงเฟรมที่ 70 จะมีการหมุนคอ และโบหน้า จากนั้นจะมีการเคลื่อนไหวบริเวณปากเพียงอย่างเดียว ทำให้ค่าความแตกต่างสัมบูรณ์เฉลี่ยมีค่าสูงขึ้นในช่วงแรก และลดต่ำลงในช่วงหลัง ในตอนท้ายของลำดับวิดีโอจะมีการเอียงกล้องทำให้วัตถุหายไปจากภาพ ค่าความแตกต่างสัมบูรณ์เฉลี่ยที่ได้จึงมีค่าสูง สำหรับลำดับวิดีโอ Carphone ในช่วงแรกในแต่จะมีความซับซ้อนไม่แตกต่างกันมาก จะมีเพียงบริเวณโบหน้าที่เคลื่อนไหว ถึงแม้จากหลังจะมีการเปลี่ยนแปลงตลอดแต่จากหลังจะเป็นท้องฟ้า ซึ่งมีค่าสีอยู่ในโทนเดียวกัน จนกระทั่งในเฟรมที่ 60 บริเวณจากหลังจะเปลี่ยนจากท้องฟ้าเป็นพื้นป่าที่มีต้นไม้ และมีการเคลื่อนไหวของโบหน้ากับมือ ทำให้ค่าความแตกต่างสัมบูรณ์เฉลี่ยมีค่าสูงขึ้น จนกระทั่งในช่วงหลังจากหลังจะเป็นพื้นป่าที่มีค่าสีอยู่ในโทนเดียวกัน และมีการเคลื่อนไหวของโบหน้าเพียงอย่างเดียว ค่าความแตกต่างสัมบูรณ์เฉลี่ยจึงมีค่าลดลง สำหรับลำดับวิดีโอ Suzie+Trevor ในช่วงเฟรมที่ 20 จะมีการเคลื่อนไหวของวัตถุออกจากเฟรมภาพจากนั้นจะเคลื่อนกลับเข้ามาใหม่ ทำให้ค่าความแตกต่างสัมบูรณ์เฉลี่ยมีค่าสูงขึ้น และลดต่ำลงในช่วงหลัง ในเฟรมที่ 50 ซึ่งเป็นการ

เปลี่ยนของลำดับวิดีโอที่สุ่มจาก Suzie เป็น Trevor ทำให้ค่าความแตกต่างสัมบูรณ์เฉลี่ยมีค่าสูงขึ้นอย่างรวดเร็ว ในช่วงแรกของลำดับภาพ Trevor จะมีวัตถุภายในเฟรมเดียวกันหลายวัตถุซึ่งเคลื่อนที่ต่างกันออกไป ทำให้ค่าความแตกต่างสัมบูรณ์เฉลี่ยมีค่าสูงขึ้น จนกระทั่งในเฟรมที่ 70 มีการตัดภาพ ทำให้ค่าความแตกต่างสัมบูรณ์เฉลี่ยมีค่าสูงที่สุด ซึ่งในตอนหลังของลำดับวิดีโอจะเหลือเพียงวัตถุเดียวที่เคลื่อนไหว ทำให้ค่าความแตกต่างสัมบูรณ์เฉลี่ยมีค่าลดลง

4.2.2 ผลการทดสอบเมื่อไม่มีข้อกำหนดของขนาดบัพเฟอรัตัวถอดรหัส และเวลาบรจูก่อน

ในหัวข้อนี้เป็นการเปรียบเทียบคุณภาพของสัญญาณวิดีโอของวิธีที่นำเสนอเมื่อไม่มีข้อกำหนดของขนาดบัพเฟอรัตัวถอดรหัส และเวลาบรจูก่อน ที่อัตราบิตเป้าหมายต่าง ๆ กัน คือที่ 32 64 128 และ 256 กิโลบิตต่อวินาที และที่ขนาดหน้าต่างที่ต่างกัน คือที่ $W = 50$ และ 100 กับการเข้ารหัสสัญญาณวิดีโอมาตรฐาน H.264 และทำการหาค่าขนาดบัพเฟอรัตัวถอดรหัสที่ต้องการ และค่าเวลาบรจูก่อนที่ต้องการ สำหรับแต่ละลำดับสัญญาณวิดีโอ

ผลการเปรียบเทียบคุณภาพของสัญญาณวิดีโอในลำดับวิดีโอต่าง ๆ จะแบ่งเป็นผลการเปรียบเทียบคุณภาพเชิงอัตวิสัย (Subjective Test) ซึ่งเป็นการเปรียบเทียบคุณภาพด้วยการมองด้วยตา โดยมีผู้ประเมินคือผู้ทำวิจัย และผลการเปรียบเทียบคุณภาพเชิงปรวิสัยด้วยค่า PSNR (Objective Test) แสดงดังรูปที่ 4.2 ถึง 4.37 และตารางที่ 4.2 ถึง 4.4 จะเป็นตารางที่สรุปรวมผลการเปรียบเทียบรวมถึงขนาดของบัพเฟอรัตัวถอดรหัส และเวลาบรจูก่อนที่ต้องการ

4.2.2.1 ผลการทดสอบกับลำดับสัญญาณวิดีโอ Foreman

ก) อัตราบิตเป้าหมาย 32 กิโลบิตต่อวินาที และขนาดหน้าต่าง $W = 50$ และ 100

ผลการเปรียบเทียบคุณภาพของสัญญาณวิดีโอของลำดับวิดีโอ Foreman ที่อัตราบิตเป้าหมาย 32 กิโลบิตต่อวินาที ที่ขนาดหน้าต่าง $W = 50$ และ 100 จะแบ่งเป็นผลการเปรียบเทียบคุณภาพเชิงอัตวิสัย จะเป็นผลการเปรียบเทียบระหว่างเฟรมที่มีค่าคุณภาพเพิ่มมากที่สุด และเฟรมที่มีค่าคุณภาพลดลงมากที่สุดของวิธีที่นำเสนอกับการเข้ารหัสมาตรฐาน H.264 เพื่อแสดงให้เห็นว่าในเฟรมที่มีคุณภาพลดลงมากที่สุดเมื่อมองด้วยตาแล้วจะสังเกตเห็นความแตกต่างได้ยาก

และในเฟรมที่มีคุณภาพเพิ่มขึ้นมากที่สุดจะสังเกตเห็นความแตกต่างได้อย่างชัดเจน แสดงดังรูปที่ 4.2 และ 4.3 และผลการเปรียบเทียบคุณภาพเชิงประจักษ์ด้วยค่า PSNR แสดงดังรูปที่ 4.4

เฟรมที่มีคุณภาพลดลงมากที่สุด

JM6.1



(ก)

ระเบียบวิธีที่นำเสนอ



(ข)

เฟรมที่มีคุณภาพเพิ่มมากที่สุด

JM6.1



(ค)

ระเบียบวิธีที่นำเสนอ



(ง)

รูปที่ 4.2 การวัดคุณภาพเชิงอัตวิสัยสำหรับลำดับวิดีโอ Foreman ที่อัตราบิตเป้าหมาย 32 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W = 50$ เฟรมที่ 73 มีคุณภาพลดลงมากที่สุด = 0.93 dB สำหรับ JM6.1 แสดงใน (ก) และระเบียบวิธีที่นำเสนอ แสดงใน (ข) เฟรมที่ 98 มีคุณภาพเพิ่มมากขึ้นที่สุด = 1.29 dB สำหรับ JM6.1 แสดงใน (ค) และระเบียบวิธีที่นำเสนอ แสดงใน (ง)

เฟรมที่มีคุณภาพลดลงมากที่สุด

JM6.1



(ก)

ระเบียบวิธีที่นำเสนอ



(ข)

เฟรมที่มีคุณภาพเพิ่มมากที่สุด

JM6.1



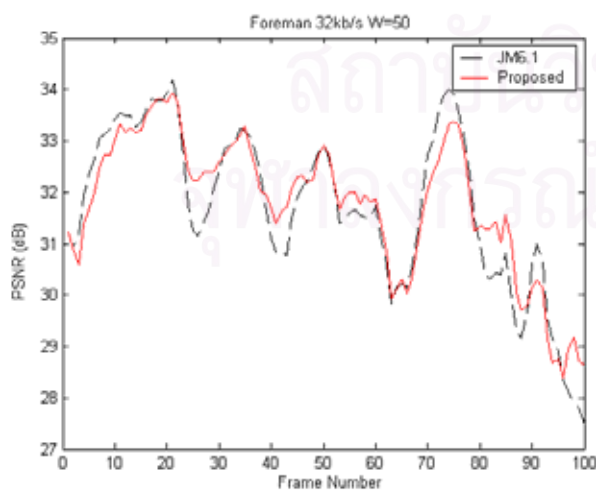
(ค)

ระเบียบวิธีที่นำเสนอ

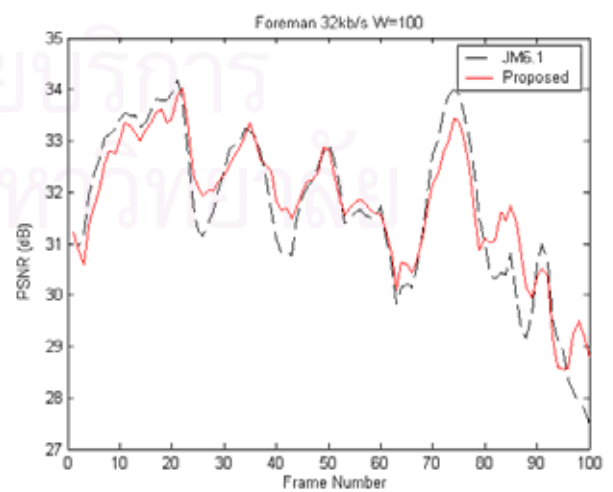


(ง)

รูปที่ 4.3 การวัดคุณภาพเชิงอัตตวิสัยสำหรับลำดับวิดีโอ Foreman ที่อัตราบิตเป้าหมาย 32 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W=100$ เฟรมที่ 73 มีคุณภาพลดลงมากที่สุด $=0.90$ dB สำหรับ JM6.1 แสดงใน (ก) และระเบียบวิธีที่นำเสนอ แสดงใน (ข) เฟรมที่ 98 มีคุณภาพเพิ่มมากขึ้นที่สุด $=1.6$ dB สำหรับ JM6.1 แสดงใน (ค) และระเบียบวิธีที่นำเสนอ แสดงใน (ง)



(ก)



(ข)

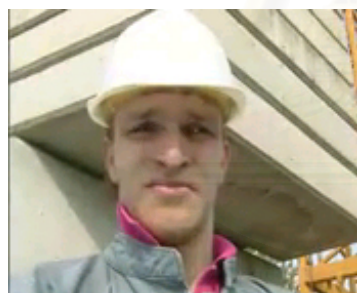
รูปที่ 4.4 กราฟการเปรียบเทียบคุณภาพลำดับวิดีโอ Foreman ที่อัตราบิตเป้าหมาย 32 กิโลบิตต่อวินาที ขนาดหน้าต่างต่าง $W=50$ แสดงใน (ก) และ ขนาดหน้าต่างต่าง $W=100$ แสดงใน (ข)

ข) อัตราบิตเป้าหมาย 64 กิโลบิตต่อวินาที และขนาดหน้าต่างต่าง $W=50$ และ 100

ผลการเปรียบเทียบคุณภาพของสัญญาณวิดีโอของลำดับวิดีโอ Foreman ที่อัตราบิตเป้าหมาย 64 กิโลบิตต่อวินาที ที่ขนาดหน้าต่างต่าง $W=50$ และ 100 จะแบ่งเป็นผลการเปรียบเทียบคุณภาพเชิงอัตตวิสัย จะเป็นผลการเปรียบเทียบระหว่างเฟรมที่มีค่าคุณภาพเพิ่มมากที่สุด และเฟรมที่มีค่าคุณภาพลดลงมากที่สุดของวิธีที่นำเสนอกับการเข้ารหัสมาตรฐาน H.264 เพื่อแสดงให้เห็นว่าในเฟรมที่มีคุณภาพลดลงมากที่สุดเมื่อมองด้วยตาแล้วจะสังเกตเห็นความแตกต่างได้ยาก และในเฟรมที่มีคุณภาพเพิ่มขึ้นมากที่สุดจะสังเกตเห็นความแตกต่างได้อย่างชัดเจน แสดงดังรูปที่ 4.5 และ 4.6 และผลการเปรียบเทียบคุณภาพเชิงปริวิสัยด้วยค่า PSNR แสดงดังรูปที่ 4.7

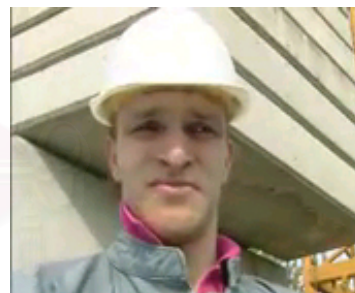
เฟรมที่มีคุณภาพลดลงมากที่สุด

JM6.1



(ก)

ระเบียบวิธีที่นำเสนอ



(ข)

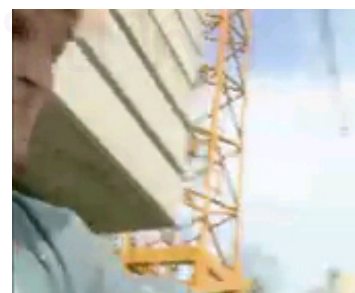
เฟรมที่มีคุณภาพเพิ่มมากที่สุด

JM6.1



(ค)

ระเบียบวิธีที่นำเสนอ

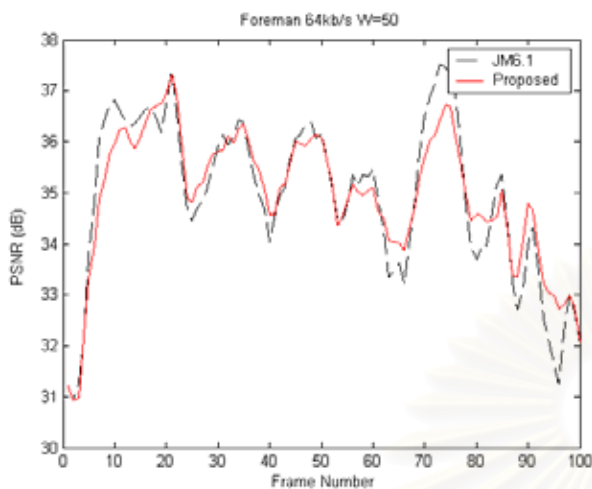


(ง)

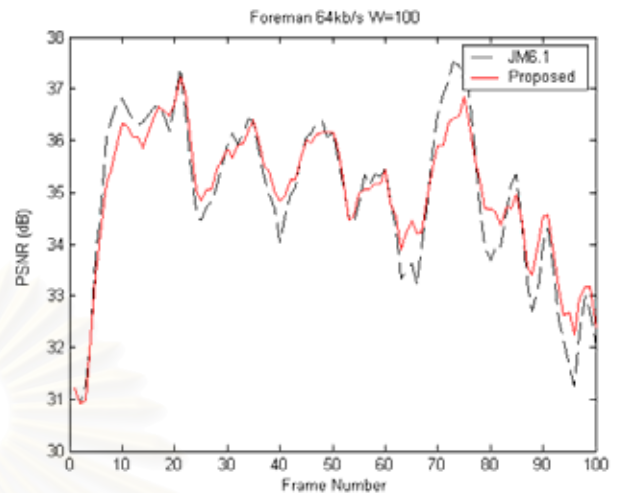
รูปที่ 4.5 การวัดคุณภาพเชิงอัตวิสัยสำหรับลำดับวิดีโอ Foreman ที่อัตราบิตเป้าหมาย 64 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W=50$ เฟรมที่ 8 มีคุณภาพลดลงมากที่สุด =1.19 dB สำหรับ JM6.1 แสดงใน (ก) และระเบียบวิธีที่นำเสนอ แสดงใน (ข) เฟรมที่ 96 มีคุณภาพเพิ่มมากขึ้นที่สุด =1.48 db สำหรับ JM6.1 แสดงใน (ค) และระเบียบวิธีที่นำเสนอ แสดงใน (ง)



รูปที่ 4.6 การวัดคุณภาพเชิงอัตวิสัยสำหรับลำดับวิดีโอ Foreman ที่อัตราบิตเป้าหมาย 64 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W=100$ เฟรมที่ 73 มีคุณภาพลดลงมากที่สุด =1.11 dB สำหรับ JM6.1 แสดงใน (ก) และระเบียบวิธีที่นำเสนอ แสดงใน (ข) เฟรมที่ 95 มีคุณภาพเพิ่มมากขึ้นที่สุด =1.06 dB สำหรับ JM6.1 แสดงใน (ค) และระเบียบวิธีที่นำเสนอ แสดงใน (ง)



(ก)



(ข)

รูปที่ 4.7 กราฟการเปรียบเทียบคุณภาพลำดับวิดีโอ Foreman ที่อัตราบิตเป้าหมาย 64 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W = 50$ แสดงใน (ก) และ ขนาดหน้าต่าง $W = 100$ แสดงใน (ข)

ค) อัตราบิตเป้าหมาย 128 กิโลบิตต่อวินาที และขนาดหน้าต่าง $W = 50$ และ 100

ผลการเปรียบเทียบคุณภาพของสัญญาณวิดีโอของลำดับวิดีโอ Foreman ที่อัตราบิตเป้าหมาย 128 กิโลบิตต่อวินาที ที่ขนาดหน้าต่าง $W = 50$ และ 100 จะแบ่งเป็นผลการเปรียบเทียบคุณภาพเชิงอัตวิสัย จะเป็นผลการเปรียบเทียบระหว่างเฟรมที่มีค่าคุณภาพเพิ่มมากที่สุด และเฟรมที่มีค่าคุณภาพลดลงมากที่สุดของวิธีที่นำเสนอกับการเข้ารหัสมาตรฐาน H.264 เพื่อแสดงให้เห็นว่าในเฟรมที่มีคุณภาพลดลงมากที่สุดเมื่อมองด้วยตาแล้วจะสังเกตเห็นความแตกต่างได้ยาก และในเฟรมที่มีคุณภาพเพิ่มขึ้นมากที่สุดจะสังเกตเห็นความแตกต่างได้อย่างชัดเจน แสดงดังรูปที่ 4.8 และ 4.9 และผลการเปรียบเทียบคุณภาพเชิงปริวิสัยด้วยค่า PSNR แสดงดังรูปที่ 4.10

จุฬาลงกรณ์มหาวิทยาลัย

เฟรมที่มีคุณภาพลดลงมากที่สุด

JM6.1



(ก)

ระเบียบวิธีที่นำเสนอ



(ข)

เฟรมที่มีคุณภาพเพิ่มมากที่สุด

JM6.1



(ค)

ระเบียบวิธีที่นำเสนอ



(ง)

รูปที่ 4.8 การวัดคุณภาพเชิงอัตวิสัยสำหรับลำดับวิดีโอ Foreman ที่อัตราบิตเป้าหมาย 128 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W=50$ เฟรมที่ 99 มีคุณภาพลดลงมากที่สุด $=1.6$ dB สำหรับ JM6.1 แสดงใน (ก) และระเบียบวิธีที่นำเสนอ แสดงใน (ข) เฟรมที่ 24 มีคุณภาพเพิ่มมากที่สุด $=0.9$ สำหรับ JM6.1 แสดงใน (ค) และระเบียบวิธีที่นำเสนอ แสดงใน (ง)

เฟรมที่มีคุณภาพลดลงมากที่สุด

JM6.1



(ก)

ระเบียบวิธีที่นำเสนอ



(ข)

เฟรมที่มีคุณภาพเพิ่มมากที่สุด

JM6.1



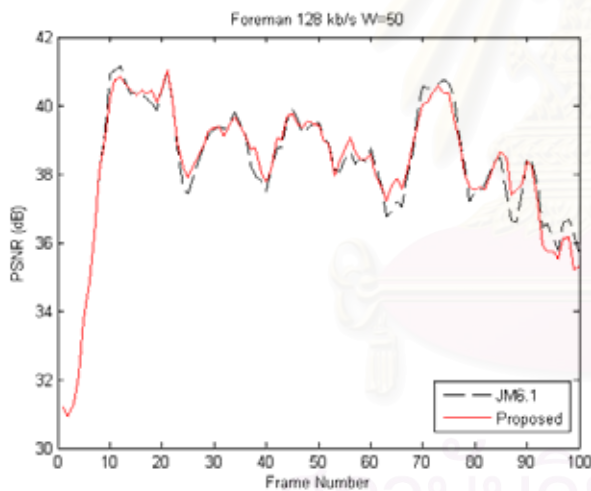
(ค)

ระเบียบวิธีที่นำเสนอ

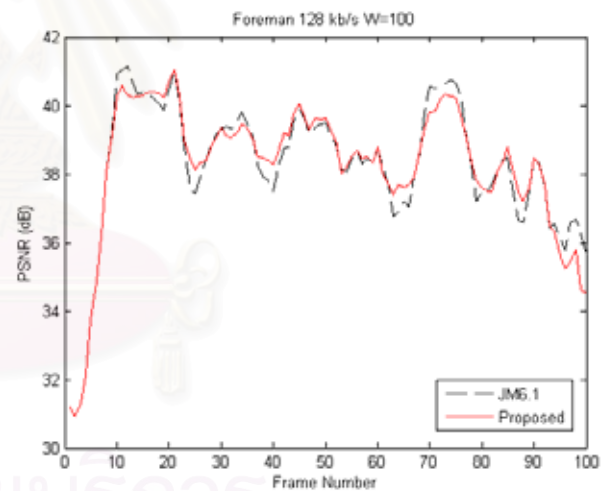


(ง)

รูปที่ 4.9 การวัดคุณภาพเชิงอัตตวิสัยสำหรับลำดับวิดีโอ Foreman ที่อัตราบิตเป้าหมาย 128 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W = 100$ เฟรมที่ 99 มีคุณภาพลดลงมากที่สุด $= 1.03$ dB สำหรับ JM6.1 แสดงใน (ก) และระเบียบวิธีที่นำเสนอ แสดงใน (ข) เฟรมที่ 86 มีคุณภาพเพิ่มมากที่สุด $= 1.06$ dB สำหรับ JM6.1 แสดงใน (ค) และระเบียบวิธีที่นำเสนอ แสดงใน (ง)



(ก)



(ข)

รูปที่ 4.10 กราฟการเปรียบเทียบคุณภาพลำดับวิดีโอ Foreman ที่อัตราบิตเป้าหมาย 128 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W = 50$ แสดงใน (ก) และ ขนาดหน้าต่าง $W = 100$ แสดงใน (ข)

ง) อัตราบิตเป้าหมาย 256 กิโลบิตต่อวินาที และขนาดหน้าต่าง $W = 50$ และ 100

ผลการเปรียบเทียบคุณภาพของสัญญาณวิดีโอของลำดับวิดีโอ Foreman ที่อัตราบิตเป้าหมาย 256 กิโลบิตต่อวินาที ที่ขนาดหน้าต่าง $W = 50$ และ 100 จะแบ่งเป็นผลการเปรียบเทียบคุณภาพเชิงอัตตวิสัย จะเป็นผลการเปรียบเทียบระหว่างเฟรมที่มีค่าคุณภาพเพิ่มมากที่สุด และ

เฟรมที่มีค่าคุณภาพลดลงมากที่สุดของวิธีที่นำเสนอกับการเข้ารหัสมาตรฐาน H.264 เพื่อแสดงให้เห็นว่าในเฟรมที่มีคุณภาพลดลงมากที่สุดเมื่อมองด้วยตาแล้วจะสังเกตเห็นความแตกต่างได้ยาก และในเฟรมที่มีคุณภาพเพิ่มขึ้นมากที่สุดจะสังเกตเห็นความแตกต่างได้อย่างชัดเจน แสดงดังรูปที่ 4.11 และ 4.12 และผลการเปรียบเทียบคุณภาพเชิงประจักษ์ด้วยค่า PSNR แสดงดังรูปที่ 4.13 และตารางที่ 4.2 จะเป็นตารางสรุปผลการเปรียบเทียบค่าคุณภาพ รวมถึงค่าขนาดของบัพเฟอร์ตัวถอดรหัส และค่าเวลาบรรจุก่อนที่ความต้องการของลำดับวิดีโอ Foreman ที่อัตราบิตเป้าหมายทั้งหมดที่ทำการทดสอบ

เฟรมที่มีคุณภาพลดลงมากที่สุด

JM6.1

ระเบียบวิธีที่นำเสนอ



(ก)



(ข)

เฟรมที่มีคุณภาพเพิ่มมากที่สุด

JM6.1

ระเบียบวิธีที่นำเสนอ



(ค)



(ง)

รูปที่ 4.11 การวัดคุณภาพเชิงอัตตวิสัยสำหรับลำดับวิดีโอ Foreman ที่อัตราบิตเป้าหมาย 256 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W=50$ เฟรมที่ 14 มีคุณภาพลดลงมากที่สุด = 2.25 dB สำหรับ JM6.1 แสดงใน (ก) และระเบียบวิธีที่นำเสนอ แสดงใน (ข) เฟรมที่ 64 มีคุณภาพเพิ่มขึ้นที่สุด = 1.52 dB สำหรับ JM6.1 แสดงใน (ค) และระเบียบวิธีที่นำเสนอ แสดงใน (ง)

เฟรมที่มีคุณภาพลดลงมากที่สุด

JM6.1



(ก)

ระเบียบวิธีที่นำเสนอ



(ข)

เฟรมที่มีคุณภาพเพิ่มมากที่สุด

JM6.1



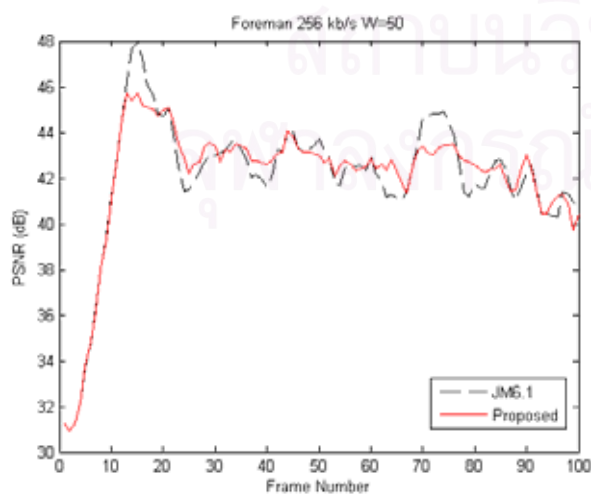
(ค)

ระเบียบวิธีที่นำเสนอ

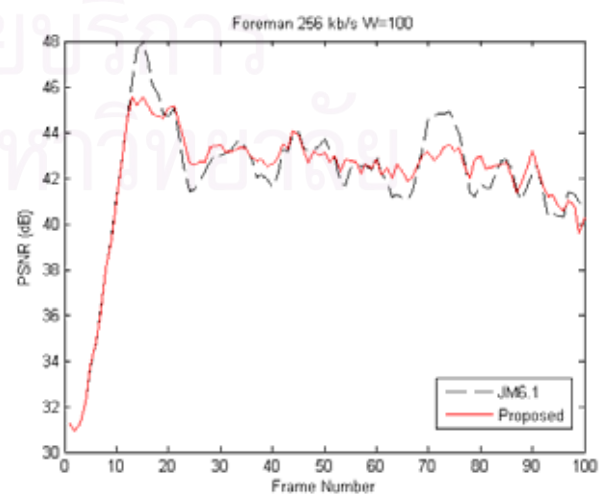


(ง)

รูปที่ 4.12 การวัดคุณภาพเชิงอัตวิสัยสำหรับลำดับวิดีโอ Foreman ที่อัตราบิตเป้าหมาย 256 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W=100$ เฟรมที่ 14 มีคุณภาพลดลงมากที่สุด $=2.47$ dB สำหรับ JM6.1 แสดงใน (ก) และระเบียบวิธีที่นำเสนอ แสดงใน (ข) เฟรมที่ 79 มีคุณภาพเพิ่มมากขึ้นที่สุด $=1.64$ dB สำหรับ JM6.1 แสดงใน (ค) และระเบียบวิธีที่นำเสนอ แสดงใน (ง)



(ก)



(ข)

รูปที่ 4.13 กราฟการเปรียบเทียบคุณภาพลำดับวิดีโอ Foreman ที่อัตราบิตเป้าหมาย 256 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W=50$ แสดงใน (ก) และ ขนาดหน้าต่าง $W=100$ แสดงใน (ข)

ตารางที่ 4.2 ผลการเปรียบเทียบคุณภาพของลำดับวิดีโอ Foreman สำหรับการเข้ารหัส JM6.1 และระเบียบวิธีที่นำเสนอ ที่ขนาดหน้าต่าง $W=50$ และ 100

Video sequence	Target Bit rate		Average PSNR (dB)	PSNR standard deviation	Average PSNR improvement (dB)	Maximum PSNR improvement (dB)	Maximum PSNR degradation (dB)	Require buffer size (kbits)	Require Pre-loading time (second)	Bit rate (kbps)
Foreman	32 kb/s	JM6.1	31.72	1.57	—	—	—	12.42	0.25	32.04
		W=50	31.79	1.35	0.07	1.29	0.93	11.22	0.19	32.44
		W=100	31.80	1.26	0.08	1.6	0.90	11.58	0.21	32.48
	64 kb/s	JM6.1	34.95	1.58	—	—	—	21.70	0.25	64.04
		W=50	34.96	1.33	0.01	1.48	1.19	21.50	0.23	64.66
		W=100	35.01	1.32	0.05	1.06	1.11	23.33	0.24	64.73
	128 kb/s	JM6.1	38.26	2.04	—	—	—	68.14	0.27	128.02
		W=50	38.27	2.03	0.01	0.9	1.60	83.36	0.38	128.47
		W=100	38.28	2.02	0.02	1.06	1.03	83.88	0.39	128.55
	256 kb/s	JM6.1	42.04	3.04	—	—	—	200.08	0.30	256.05
		W=50	42.06	2.87	0.02	1.52	2.25	207.57	0.33	256.63
		W=100	42.06	2.85	0.02	1.64	2.47	207.86	0.33	256.92

จากผลการทดสอบคุณภาพของลำดับวิดีโอ Foreman เซึ่งอัตตวิสัย และเชิงปรวิสัย เรา จะเห็นว่าระเบียบวิธีที่นำเสนอจะให้คุณภาพที่เพิ่มขึ้นมากกว่าวิธีมาตรฐานในหลาย ๆ เฟรม และ ในเฟรมที่มีคุณภาพลดลงมากที่สุดก็จะสังเกตเห็นได้ยากโดยเฉพาะที่อัตราบิตต่ำ (32 กิโลบิตต่อ วินาที) ในช่วงท้ายของลำดับวิดีโอ ซึ่งเป็นตอนที่มีการเคลื่อนกล้องทำให้วัตถุหายไปจากภาพ ซึ่ง มีความซับซ้อนที่สูง เทคนิคที่นำเสนอจะจัดสรรบิตให้มากทำให้คุณภาพที่ได้ดีกว่า สังเกตจาก บริเวณนิ้วมือ และแขนในเฟรมที่มีคุณภาพเพิ่มขึ้นมากที่สุด เทคนิคที่นำเสนอจะให้คุณภาพที่ ดีกว่าเมื่อเปรียบเทียบกับ JM6.1 ในเฟรมที่มีค่าความแตกต่างสัมบูรณ์เฉลี่ยสูงที่สุดจะไม่จำเป็น เสมอไปว่าเทคนิคที่นำเสนอจะให้คุณภาพสูงที่สุด เพราะจากการทดสอบบิตเป้าหมายจะยังคงมี ค่าต่างจากบิตที่ใช้ในการเข้ารหัสจริงค่อนข้างมาก และการคำนวณหาบิตเป้าหมายในระดับเฟรม จะยังคงเป็นบิตเป้าหมายที่มาจากระดับบัพเฟอร์ และบิตเป้าหมายจากเทคนิคที่นำเสนอ แต่จะมี การถ่วงน้ำหนักให้บิตเป้าหมายที่มาจากเทคนิคที่นำเสนอมากกว่าตามสมการ (2.11) บิต

เป้าหมายสุดท้ายในระดับเฟรมที่เฟรมที่มีค่าความแตกต่างสัมบูรณ์เฉลี่ยสูงสุด จึงไม่จำเป็นต้องมีค่ามากที่สุด คุณภาพในเฟรมนั้นจึงไม่จำเป็นต้องมีค่ามากที่สุดด้วยเหมือนกัน

จากตารางที่ 4.2 และจากกราฟที่แสดงค่า PSNR จะเห็นว่าระเบียบวิธีที่นำเสนอจะให้คุณภาพโดยเฉลี่ยสูงกว่าวิธีมาตรฐานในช่วง 0.01 ถึง 0.08 dB และเฟรมที่มีคุณภาพเพิ่มขึ้นมากที่สุดเพิ่มขึ้น 1.64 dB และในช่วงที่ค่า PSNR มีค่าตกลง ระเบียบวิธีที่นำเสนอจะมีการปรับปรุงให้ดีขึ้นเมื่อเปรียบเทียบกับ JM6.1 ถึงแม้ว่าที่อัตราบิตสูง (128 และ 256 กิโลบิตต่อวินาที) ในเฟรมที่มีคุณภาพเพิ่มขึ้นมากที่สุดจะเพิ่มขึ้นน้อยกว่าการลดลงของคุณภาพของเฟรมที่มีคุณภาพลดลงมากที่สุด แต่คุณภาพโดยเฉลี่ยทั้งหมดของลำดับวิดีโอที่สนเมื่อวัดด้วยค่า PSNR ยังคงมีค่าสูงกว่าเมื่อเปรียบเทียบกับ JM6.1 การจัดสรรบิตซ้ำจะทำการประหยัดบิตจากเฟรมที่มีกิจกรรมต่ำ และใช้บิตที่ได้จากการประหยัดนั้นกับเฟรมที่มีกิจกรรมสูงในการผลิตวิดีโอที่สนให้มีคุณภาพที่ราบเรียบ เห็นได้จากค่าเบี่ยงเบนมาตรฐานของระเบียบวิธีที่นำเสนอจะมีค่าน้อยกว่า

ขนาดของหน้าต่างที่มากจะให้คุณภาพของสัญญาณวิดีโอที่สนที่ราบเรียบกว่าขนาดของหน้าต่างที่มีขนาดเล็กกว่า เฟรมที่มีคุณภาพเพิ่มมากที่สุดของระเบียบวิธีที่นำเสนอที่ขนาดหน้าต่าง 100 จะมีค่ามากกว่าที่ขนาดหน้าต่าง 50 นั่นเป็นเพราะจะมีข้อมูลของเฟรมในอนาคตที่มากกว่านำมาใช้ในการจัดสรรบิตซ้ำ ทั้งนี้ค่าเบี่ยงเบนมาตรฐานก็จะมีค่าน้อยกว่าด้วย ซึ่งจะบ่งบอกถึงว่าเมื่อขนาดหน้าต่างมีขนาดใหญ่กว่า คุณภาพของสัญญาณวิดีโอที่สนจะเกาะกลุ่มกันมากกว่า เพราะจำนวนบิตจะถูกเฉลี่ยไปในช่วงที่กว้างกว่า และที่ขนาดหน้าต่างทั้ง 50 และ 100 ค่าเบี่ยงเบนมาตรฐานของค่า PSNR จะมีค่าน้อยกว่าวิธีมาตรฐานอีกด้วย ผลของขนาดหน้าต่างต่อขนาดของบัพเฟอรัต์ตัวถอดรหัส และเวลาบรจุก่อน ที่ขนาดหน้าต่าง 100 จะต้องการเวลาบรจุก่อน และขนาดของบัพเฟอรัต์ตัวถอดรหัสที่มากกว่าที่ขนาดหน้าต่าง 50 เนื่องมาจากจำนวนบิตจะถูกเฉลี่ยไปในช่วงที่กว้างกว่า ทำให้ในหลาย ๆ เฟรมจะมีจำนวนบิตที่มากกว่า ดูได้จากจำนวนอัตราบิตที่ได้ จึงทำให้ต้องการเวลาบรจุก่อน และขนาดของบัพเฟอรัต์ตัวถอดรหัสที่มากกว่า

ข้อดีของระเบียบวิธีที่นำเสนอสำหรับลำดับวิดีโอที่สน Foreman จะมีความต้องการขนาดของบัพเฟอรัต์ตัวถอดรหัส และเวลาบรจุก่อนมากกว่าเมื่อเปรียบเทียบกับ JM6.1 เนื่องมาจากเทคนิคที่นำเสนอมีการจัดสรรบิตให้เฉลี่ยไปในหลาย ๆ เฟรม ทำให้ในหลาย ๆ เฟรมจำนวนบิตที่ใช้มีค่ามากกว่าเมื่อเปรียบเทียบกับ JM6.1 ขนาดของบัพเฟอรัต์ตัวถอดรหัสที่ต้องการ และเวลาบรจุก่อนจึงมีค่ามากกว่า แต่ดังแสดงในตาราง บัพเฟอรัต์ตัวถอดรหัสที่ต้องการโดยเฉลี่ยเพียงไม่กี่ร้อย

กิโลบิต และเวลาบรรจุก่อนที่ต้องการก็มีค่าน้อยกว่า 1 วินาที ซึ่งเป็นเหตุเป็นผลสำหรับในงาน
ประยุกต์วีดิทัศน์ทั่ว ๆ ไป

4.2.2.2 ผลการทดสอบกับลำดับสัญญาณวีดิทัศน์ Carphone

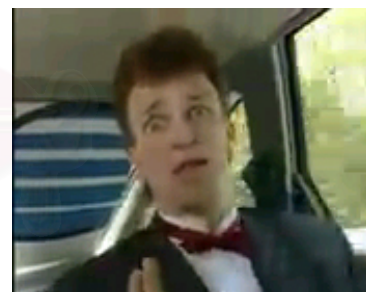
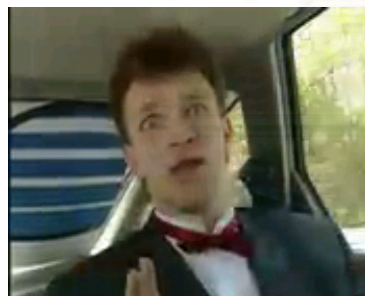
ก) อัตราบิตเป้าหมาย 32 กิโลบิตต่อวินาที และขนาดหน้าต่าง $W = 50$ และ 100

ผลการเปรียบเทียบคุณภาพของสัญญาณวีดิทัศน์ของลำดับวีดิทัศน์ Carphone ที่อัตรา
บิตเป้าหมาย 32 กิโลบิตต่อวินาที ที่ขนาดหน้าต่าง $W = 50$ และ 100 จะแบ่งเป็นผลการ
เปรียบเทียบคุณภาพเชิงอัตวิสัย จะเป็นผลการเปรียบเทียบระหว่างเฟรมที่มีค่าคุณภาพเพิ่มมาก
ที่สุด และเฟรมที่มีค่าคุณภาพลดลงมากที่สุดของวิธีที่นำเสนอกับการเข้ารหัสมาตรฐาน H.264
เพื่อแสดงให้เห็นว่าในเฟรมที่มีคุณภาพลดลงมากที่สุดเมื่อมองด้วยตาแล้วจะสังเกตเห็นความ
แตกต่างได้ยาก และในเฟรมที่มีคุณภาพเพิ่มขึ้นมากที่สุดจะสังเกตเห็นความแตกต่างได้อย่าง
ชัดเจน แสดงดังรูปที่ 4.14 และ 4.15 และผลการเปรียบเทียบคุณภาพเชิงปรวิสัยด้วยค่า PSNR
แสดงดังรูปที่ 4.16

เฟรมที่มีคุณภาพลดลงมากที่สุด

JM6.1

ระเบียบวิธีที่นำเสนอ



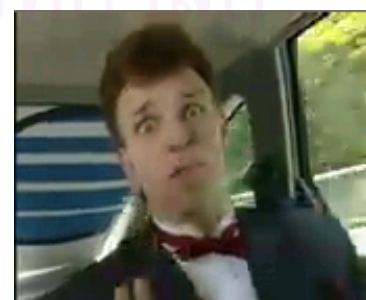
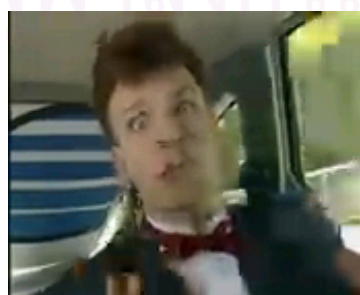
(ก)

(ข)

เฟรมที่มีคุณภาพเพิ่มมากที่สุด

JM6.1

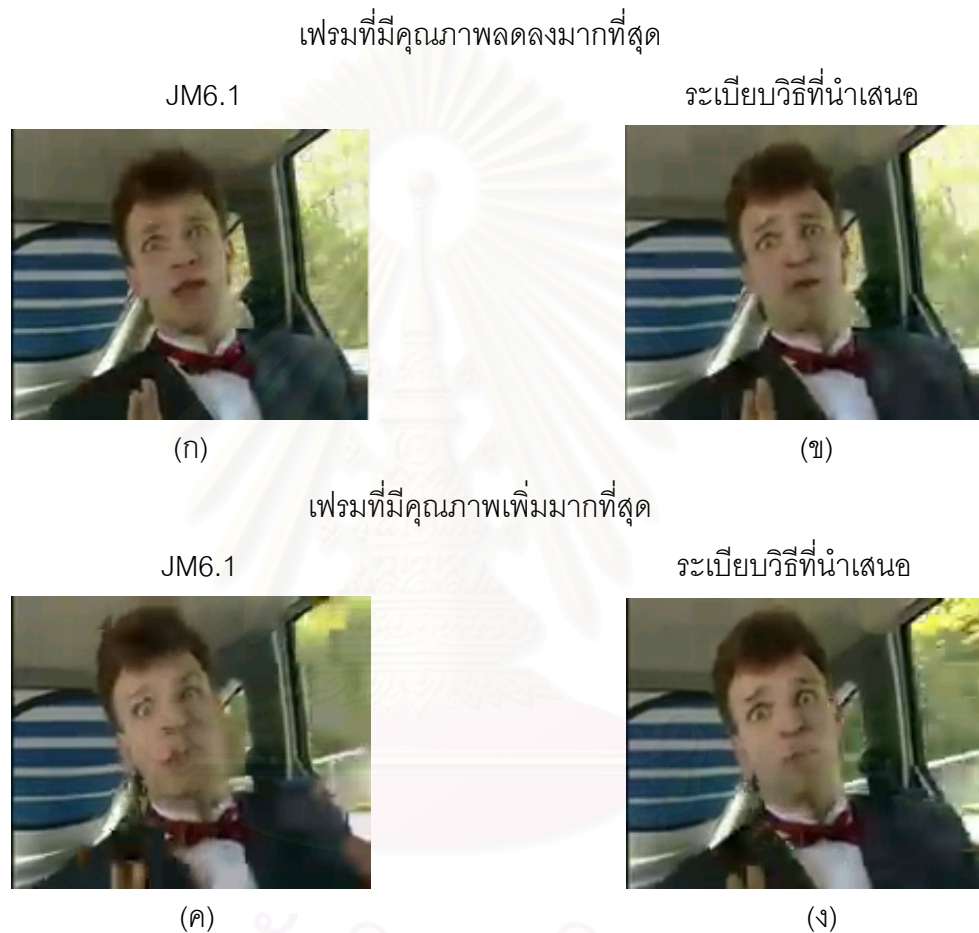
ระเบียบวิธีที่นำเสนอ



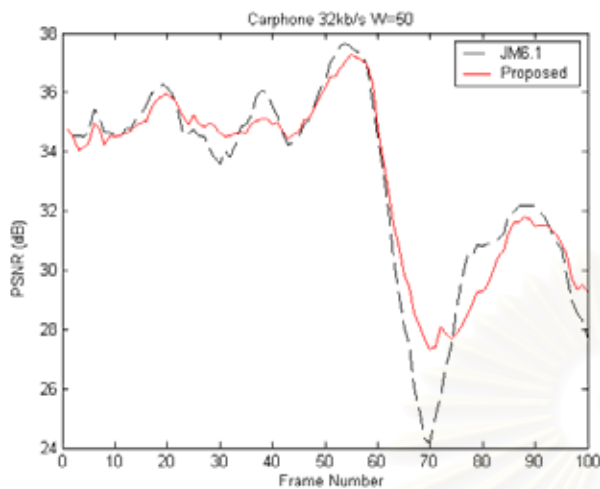
(ค)

(ง)

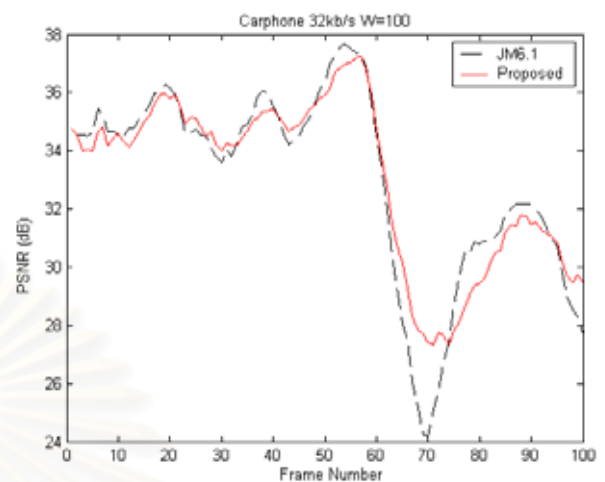
รูปที่ 4.14 การวัดคุณภาพเชิงอัตวิสัยสำหรับลำดับวิดีโอ Carphone ที่อัตราบิตเป้าหมาย 32 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W=50$ เฟรมที่ 77 มีคุณภาพลดลงมากที่สุด = 1.98 dB สำหรับ JM6.1 แสดงใน (ก) และระเบียบวิธีที่นำเสนอ แสดงใน (ข) เฟรมที่ 69 มีคุณภาพเพิ่มมากขึ้นที่สุด = 3.37 dB สำหรับ JM6.1 แสดงใน (ค) และระเบียบวิธีที่นำเสนอ แสดงใน (ง)



รูปที่ 4.15 การวัดคุณภาพเชิงอัตวิสัยสำหรับลำดับวิดีโอ Carphone ที่อัตราบิตเป้าหมาย 32 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W=100$ เฟรมที่ 77 มีคุณภาพลดลงมากที่สุด = 1.88 dB สำหรับ JM6.1 แสดงใน (ก) และระเบียบวิธีที่นำเสนอ แสดงใน (ข) เฟรมที่ 69 มีคุณภาพเพิ่มมากขึ้นที่สุด = 3.42 dB สำหรับ JM6.1 แสดงใน (ค) และระเบียบวิธีที่นำเสนอ แสดงใน (ง)



(ก)



(ข)

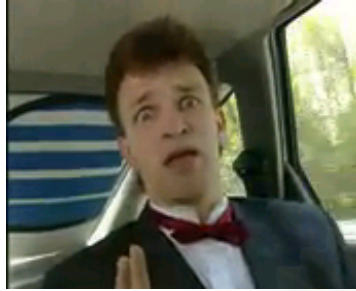
รูปที่ 4.16 กราฟการเปรียบเทียบคุณภาพลำดับวีดิทัศน์ Carphone ที่อัตราบิตเป้าหมาย 32 กิโลบิตต่อวินาที ขนาดหน้าต่างต่าง $W = 50$ แสดงใน (ก) และ ขนาดหน้าต่างต่าง $W = 100$ แสดงใน (ข)

ข) อัตราบิตเป้าหมาย 64 กิโลบิตต่อวินาที และขนาดหน้าต่าง $W = 50$ และ 100

ผลการเปรียบเทียบคุณภาพของสัญญาณวีดิทัศน์ของลำดับวีดิทัศน์ Carphone ที่อัตราบิตเป้าหมาย 64 กิโลบิตต่อวินาที ที่ขนาดหน้าต่างต่าง $W = 50$ และ 100 จะแบ่งเป็นผลการเปรียบเทียบคุณภาพเชิงอัตวิสัย จะเป็นผลการเปรียบเทียบระหว่างเฟรมที่มีค่าคุณภาพเพิ่มมากที่สุด และเฟรมที่มีค่าคุณภาพลดลงมากที่สุดของวิธีที่นำเสนอกับการเข้ารหัสมาตรฐาน H.264 เพื่อแสดงให้เห็นว่าในเฟรมที่มีคุณภาพลดลงมากที่สุดเมื่อมองด้วยตาแล้วจะสังเกตเห็นความแตกต่างได้ยาก และในเฟรมที่มีคุณภาพเพิ่มขึ้นมากที่สุดจะสังเกตเห็นความแตกต่างได้อย่างชัดเจน แสดงดังรูปที่ 4.17 และ 4.18 และผลการเปรียบเทียบคุณภาพเชิงปริวิสัยด้วยค่า PSNR แสดงดังรูปที่ 4.19

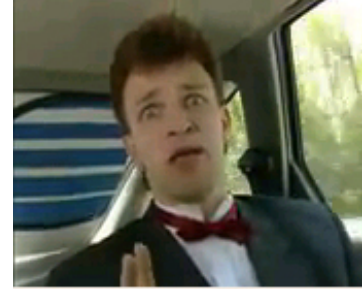
เฟรมที่มีคุณภาพลดลงมากที่สุด

JM6.1



(ก)

ระเบียบวิธีที่นำเสนอ



(ข)

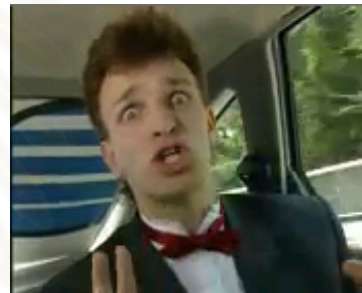
เฟรมที่มีคุณภาพเพิ่มมากที่สุด

JM6.1



(ค)

ระเบียบวิธีที่นำเสนอ

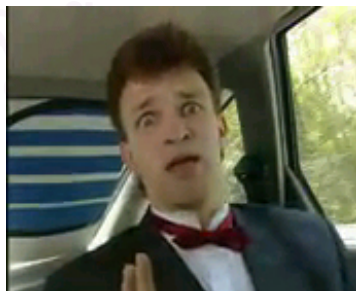


(ง)

รูปที่ 4.17 การวัดคุณภาพเชิงอัตตวิสัยสำหรับลำดับวิดีโอ Carphone ที่อัตราบิตเป้าหมาย 64 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W = 50$ เฟรมที่ 77 มีคุณภาพลดลงมากที่สุด $= 1.56$ dB สำหรับ JM6.1 แสดงใน (ก) และระเบียบวิธีที่นำเสนอ แสดงใน (ข) เฟรมที่ 68 มีคุณภาพเพิ่มมากที่สุด $= 2.75$ dB สำหรับ JM6.1 แสดงใน (ค) และระเบียบวิธีที่นำเสนอ แสดงใน (ง)

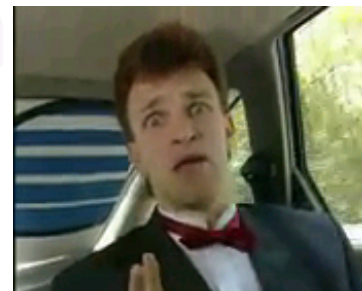
เฟรมที่มีคุณภาพลดลงมากที่สุด

JM6.1



(ก)

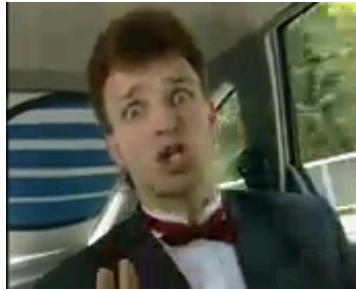
ระเบียบวิธีที่นำเสนอ



(ข)

เฟรมที่มีคุณภาพเพิ่มมากที่สุด

JM6.1



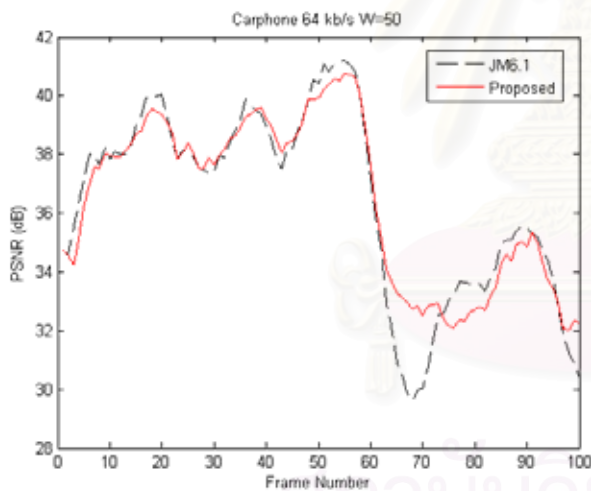
(ค)

ระเบียบวิธีที่นำเสนอ

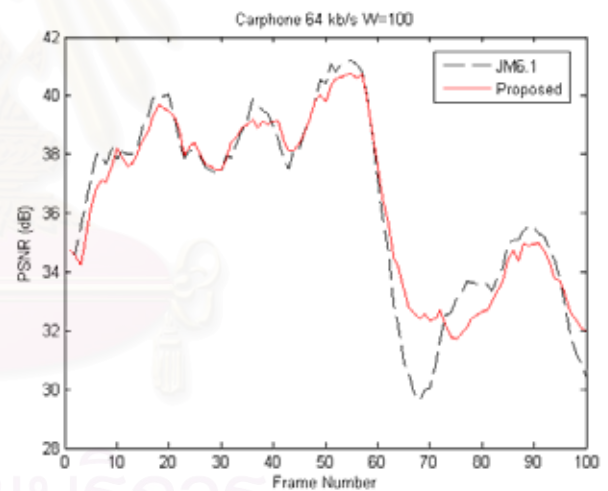


(ง)

รูปที่ 4.18 การวัดคุณภาพเชิงอัตตวิสัยสำหรับลำดับวิดีโอ Carphone ที่อัตราบิตเป้าหมาย 64 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W=100$ เฟรมที่ 77 มีคุณภาพลดลงมากที่สุด $=1.33$ dB สำหรับ JM6.1 แสดงใน (ก) และระเบียบวิธีที่นำเสนอ แสดงใน (ข) เฟรมที่ 68 มีคุณภาพเพิ่มมากขึ้นที่สุด $=3.10$ dB สำหรับ JM6.1 แสดงใน (ค) และระเบียบวิธีที่นำเสนอ แสดงใน (ง)



(ก)



(ข)

รูปที่ 4.19 กราฟการเปรียบเทียบคุณภาพลำดับวิดีโอ Carphone ที่อัตราบิตเป้าหมาย 64 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W=50$ แสดงใน (ก) และ ขนาดหน้าต่าง $W=100$ แสดงใน (ข)

ค) อัตราบิตเป้าหมาย 128 กิโลบิตต่อวินาที และขนาดหน้าต่าง $W=50$ และ 100

ผลการเปรียบเทียบคุณภาพของสัญญาณวิดีโอของลำดับวิดีโอ Carphone ที่อัตราบิตเป้าหมาย 128 กิโลบิตต่อวินาที ที่ขนาดหน้าต่าง $W=50$ และ 100 จะแบ่งเป็นผลการเปรียบเทียบคุณภาพเชิงอัตตวิสัย จะเป็นผลการเปรียบเทียบระหว่างเฟรมที่มีค่าคุณภาพเพิ่มมากที่สุด

ที่สุด และเฟรมที่มีค่าคุณภาพลดลงมากที่สุดของวิธีที่นำเสนอกับการเข้ารหัสมาตรฐาน H.264 เพื่อแสดงให้เห็นว่าในเฟรมที่มีคุณภาพลดลงมากที่สุดเมื่อมองด้วยตาแล้วจะสังเกตเห็นความแตกต่างได้ยาก และในเฟรมที่มีคุณภาพเพิ่มขึ้นมากที่สุดจะสังเกตเห็นความแตกต่างได้อย่างชัดเจน แสดงดังรูปที่ 4.20 และ 4.21 และผลการเปรียบเทียบคุณภาพเชิงประจักษ์ด้วยค่า PSNR แสดงดังรูปที่ 4.22

เฟรมที่มีคุณภาพลดลงมากที่สุด

JM6.1



(ก)

ระเบียบวิธีที่นำเสนอ



(ข)

เฟรมที่มีคุณภาพเพิ่มมากที่สุด

JM6.1



(ค)

ระเบียบวิธีที่นำเสนอ



(ง)

รูปที่ 4.20 การวัดคุณภาพเชิงอัตวิสัยสำหรับลำดับวิดีโอ Carphone ที่อัตราบิตเป้าหมาย 128 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W = 50$ เฟรมที่ 79 มีคุณภาพลดลงมากที่สุด $= 2.13$ dB สำหรับ JM6.1 แสดงใน (ก) และระเบียบวิธีที่นำเสนอ แสดงใน (ข) เฟรมที่ 63 มีคุณภาพเพิ่มมากที่สุด $= 3.47$ dB สำหรับ JM6.1 แสดงใน (ค) และระเบียบวิธีที่นำเสนอ แสดงใน (ง)

เฟรมที่มีคุณภาพลดลงมากที่สุด

JM6.1



(ก)

ระเบียบวิธีที่นำเสนอ



(ข)

เฟรมที่มีคุณภาพเพิ่มมากที่สุด

JM6.1



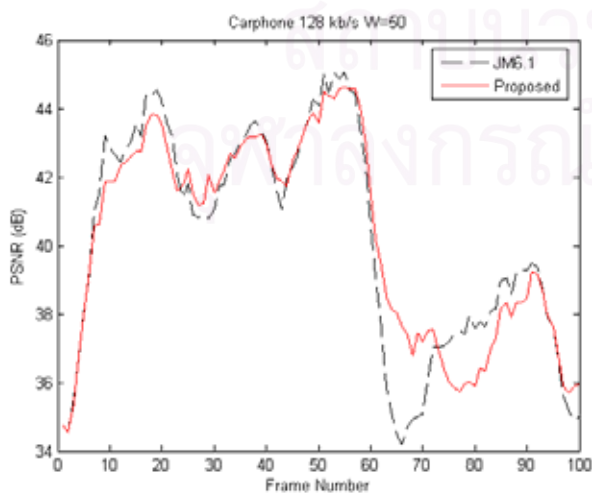
(ค)

ระเบียบวิธีที่นำเสนอ

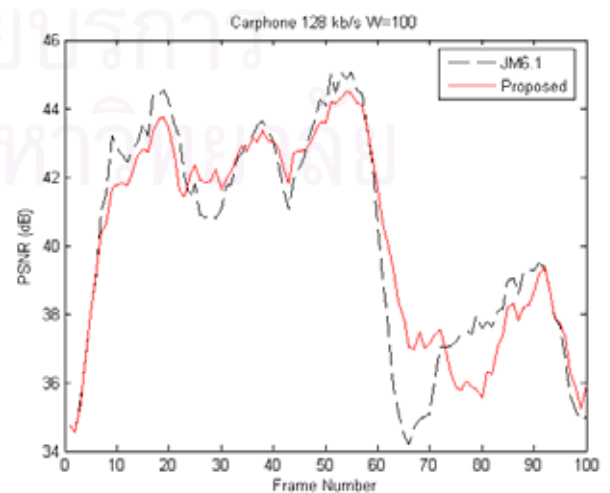


(ง)

รูปที่ 4.21 การวัดคุณภาพเชิงอัตวิสัยสำหรับลำดับวิดีโอ Carphone ที่อัตราบิตเป้าหมาย 128 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W=100$ เฟรมที่ 79 มีคุณภาพลดลงมากที่สุด $=1.92$ dB สำหรับ JM6.1 แสดงใน (ก) และระเบียบวิธีที่นำเสนอ แสดงใน (ข) เฟรมที่ 65 มีคุณภาพเพิ่มมากขึ้นที่สุด $=3.60$ dB สำหรับ JM6.1 แสดงใน (ค) และระเบียบวิธีที่นำเสนอ แสดงใน (ง)



(ก)



(ข)

รูปที่ 4.22 กราฟการเปรียบเทียบคุณภาพลำดับวิดีโอ Carphone ที่อัตราบิตเป้าหมาย 128 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W=50$ แสดงใน (ก) และ ขนาดหน้าต่าง $W=100$ แสดงใน (ข)

ง) อัตราบิตเป้าหมาย 256 กิโลบิตต่อวินาที และขนาดหน้าต่าง $W=50$ และ 100

ผลการเปรียบเทียบคุณภาพของสัญญาณวิดีโอของลำดับวิดีโอ Carphone ที่อัตราบิตเป้าหมาย 256 กิโลบิตต่อวินาที ที่ขนาดหน้าต่าง $W=50$ และ 100 จะแบ่งเป็นผลการเปรียบเทียบคุณภาพเชิงอัตวิสัย จะเป็นผลการเปรียบเทียบระหว่างเฟรมที่มีค่าคุณภาพเพิ่มมากที่สุด และเฟรมที่มีค่าคุณภาพลดลงมากที่สุดของวิธีที่นำเสนอกับการเข้ารหัสมาตรฐาน H.264 เพื่อแสดงให้เห็นว่าในเฟรมที่มีคุณภาพลดลงมากที่สุดเมื่อมองด้วยตาแล้วจะสังเกตเห็นความแตกต่างได้ยาก และในเฟรมที่มีคุณภาพเพิ่มขึ้นมากที่สุดจะสังเกตเห็นความแตกต่างได้อย่างชัดเจน แสดงดังรูปที่ 4.23 และ 4.24 และผลการเปรียบเทียบคุณภาพเชิงปริวิสัยด้วยค่า PSNR แสดงดังรูปที่ 4.25 และตารางที่ 4.3 จะเป็นตารางสรุปผลการเปรียบเทียบค่าคุณภาพ รวมถึงค่าขนาดของบัพเฟอร์ตัวถอดรหัส และค่าเวลาบรจุก่อนที่ต่อการของลำดับวิดีโอ Carphone ที่อัตราบิตเป้าหมายทั้งหมดที่ทำการทดสอบ

เฟรมที่มีคุณภาพลดลงมากที่สุด

JM6.1

ระเบียบวิธีที่นำเสนอ



(ก)



(ข)

เฟรมที่มีคุณภาพเพิ่มมากที่สุด

JM6.1



(ค)

ระเบียบวิธีที่นำเสนอ



(ง)

รูปที่ 4.23 การวัดคุณภาพเชิงอัตวิสัยสำหรับลำดับวิดีโอ Carphone ที่อัตราบิตเป้าหมาย 256 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W = 50$ เฟรมที่ 79 มีคุณภาพลดลงมากที่สุด $= 3.07$ dB สำหรับ JM6.1 แสดงใน (ก) และระเบียบวิธีที่นำเสนอ แสดงใน (ข) เฟรมที่ 66 มีคุณภาพเพิ่มมากที่สุด $= 4.98$ dB สำหรับ JM6.1 แสดงใน (ค) และระเบียบวิธีที่นำเสนอ แสดงใน (ง)

เฟรมที่มีคุณภาพลดลงมากที่สุด

JM6.1



(ก)

ระเบียบวิธีที่นำเสนอ



(ข)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

เฟรมที่มีคุณภาพเพิ่มมากที่สุด

JM6.1



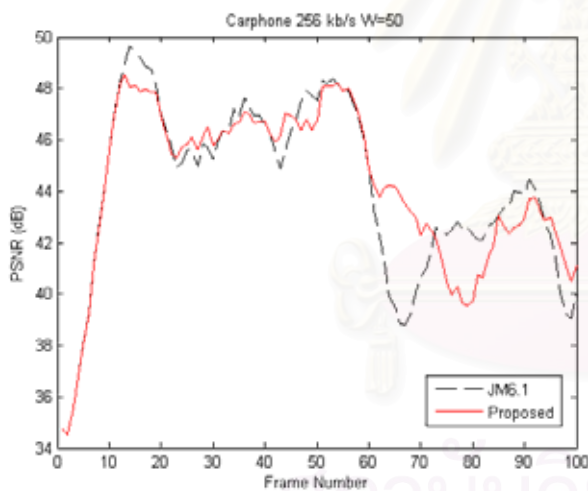
(ค)

ระเบียบวิธีที่นำเสนอ

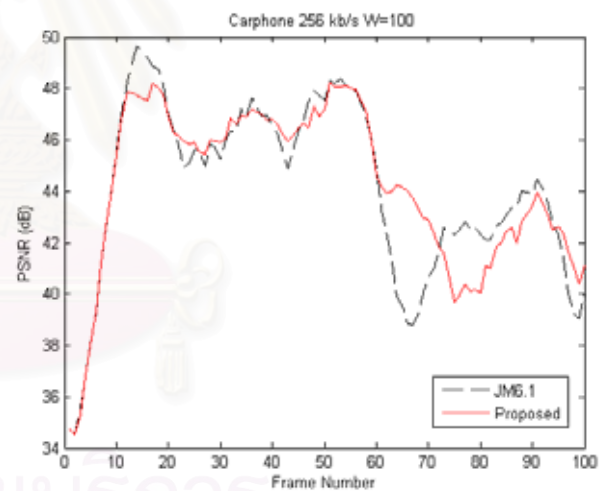


(ง)

รูปที่ 4.24 การวัดคุณภาพเชิงอัตวิสัยสำหรับลำดับวิดีโอ Carphone ที่อัตราบิตเป้าหมาย 256 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W=100$ เฟรมที่ 75 มีคุณภาพลดลงมากที่สุด $=2.61$ dB สำหรับ JM6.1 แสดงใน (ก) และระเบียบวิธีที่นำเสนอ แสดงใน (ข) เฟรมที่ 66 มีคุณภาพเพิ่มมากที่สุด $=5.14$ dB สำหรับ JM6.1 แสดงใน (ค) และระเบียบวิธีที่นำเสนอ แสดงใน (ง)



(ก)



(ข)

รูปที่ 4.25 กราฟการเปรียบเทียบคุณภาพลำดับวิดีโอ Carphone ที่อัตราบิตเป้าหมาย 256 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W=50$ แสดงใน (ก) และ ขนาดหน้าต่าง $W=100$ แสดงใน (ข)

ตารางที่ 4.3 ผลการเปรียบเทียบคุณภาพของลำดับวิดีโอ Carphone สำหรับการเข้ารหัส JM6.1 และระเบียบวิธีที่นำเสนอ ที่ขนาดหน้าต่าง $W = 50$ และ 100

Video sequence	Target Bit rate		Average PSNR (dB)	PSNR standard deviation	Average PSNR improvement (dB)	Maximum PSNR improvement (dB)	Maximum PSNR degradation (dB)	Require buffer size (kbits)	Require Pre-loading time (second)	Bit rate (kbps)
Car phone	32 kb/s	JM6.1	33.01	3.32	—	—	—	21.51	0.34	32.17
		W=50	33.11	2.86	0.10	3.37	1.98	21.27	0.39	32.89
		W=100	33.06	2.83	0.05	3.42	1.88	21.85	0.40	32.88
	64 kb/s	JM6.1	36.42	3.21	—	—	—	23.34	0.30	64.17
		W=50	36.43	2.86	0.01	2.75	1.56	36.13	0.33	64.54
		W=100	36.46	2.83	0.04	3.10	1.33	36.99	0.35	64.56
	128 kb/s	JM6.1	40.19	3.27	—	—	—	59.83	0.29	128.06
		W=50	40.20	2.99	0.01	3.47	2.13	80.64	0.45	128.32
		W=100	40.21	2.96	0.03	3.60	1.92	83.14	0.45	128.70
	256 kb/s	JM6.1	44.16	3.45	—	—	—	169.86	0.28	255.00
		W=50	44.23	3.19	0.07	4.98	3.07	234.39	0.53	256.13
		W=100	44.23	3.19	0.07	5.14	2.61	237.95	0.55	256.22

จากผลการทดสอบคุณภาพของลำดับวิดีโอ Carphone เริงอัตตวิสัย และเชิงปรวิสัย เราจะเห็นว่าระเบียบวิธีที่นำเสนอจะให้คุณภาพที่เพิ่มขึ้นมากกว่าวิธีมาตรฐานในหลาย ๆ เฟรม และในเฟรมที่มีคุณภาพลดลงมากที่สุดก็จะสังเกตเห็นได้ยากโดยเฉพาะที่อัตราบิตต่ำ (32 กิโลบิตต่อวินาที) ในช่วงเฟรมที่ 60 จนถึงเฟรมที่ 70 ของลำดับวิดีโอ ซึ่งเป็นช่วงที่มีการเปลี่ยนของฉากหลังแบบฉับพลัน และมีการเคลื่อนไหวของใบหน้า และมีมือ อย่างรวดเร็ว ซึ่งมีความซับซ้อนที่สูง ทำให้ค่า PSNR ตกลงอย่างรวดเร็ว เทคนิคที่นำเสนอจะจัดสรรบิตให้มากทำให้คุณภาพที่ได้ดีกว่าสังเกตจากบริเวณใบหน้า และฉากหลังในลำดับวิดีโอในเฟรมที่มีคุณภาพเพิ่มขึ้นมากที่สุด เทคนิคที่นำเสนอจะให้คุณภาพที่ดีกว่าเมื่อเปรียบเทียบกับ JM6.1 ในเฟรมที่มีค่าความแตกต่างสัมบูรณ์เฉลี่ยสูงสุดจะไม่จำเป็นเสมอไปว่าเทคนิคที่นำเสนอจะให้คุณภาพสูงสุด เพราะจากการทดสอบบิตเป้าหมายจะยังคงมีค่าต่างจากบิตที่ใช้ในการเข้ารหัสจริงค่อนข้างมาก และการคำนวณหาบิตเป้าหมายในระดับเฟรม จะยังคงเป็นบิตเป้าหมายที่มาจากระดับบัพเฟอร์ และบิตเป้าหมายจากเทคนิคที่นำเสนอ แต่จะมีการถ่วงน้ำหนักให้บิตเป้าหมายที่มาจากเทคนิคที่นำเสนอมากกว่าตามสมการ (2.11) บิตเป้าหมายสุดท้ายในระดับเฟรมที่มีเฟรมที่มีค่าความแตกต่างสัมบูรณ์เฉลี่ยสูงสุด จึงไม่จำเป็นต้องมีค่ามากที่สุด คุณภาพในเฟรมนั้นจึงไม่จำเป็นต้องมีค่ามากที่สุดด้วยเหมือนกัน

จากตารางที่ 4.3 และจากกราฟที่แสดงค่า PSNR จะเห็นว่าระเบียบวิธีที่นำเสนอจะให้คุณภาพโดยเฉลี่ยสูงกว่าวิธีมาตรฐานในช่วง 0.01 ถึง 0.10 dB และเฟรมที่มีคุณภาพเพิ่มขึ้นมากที่สุดเพิ่มขึ้นถึง 5.14 dB ซึ่งเป็นในช่วงที่มีการเปลี่ยนของฉากหลังที่ค่า PSNR มีค่าตกลงอย่างรวดเร็ว ระเบียบวิธีที่นำเสนอจะมีการดึงไม่ให้คุณภาพตกลงไปมากเพราะว่าตัวเข้ารหัสจะมีข้อมูลของเฟรมในอนาคต การเปลี่ยนของฉากหลังซึ่งมีผลให้เกิดความไม่ต่อเนื่องของความซับซ้อนของเนื้อหาวิดีโอทัศน์จะสามารถตรวจพบได้ล่วงหน้า คุณภาพโดยเฉลี่ยทั้งหมดของลำดับวิดีโอทัศน์เมื่อวัดด้วยค่า PSNR จะมีค่าสูงกว่า เมื่อเปรียบเทียบกับ JM6.1 การจัดสรรบิตซ้ำจะทำการประหยัดบิตจากเฟรมที่มีกิจกรรมต่ำ และใช้บิตที่ได้จากการประหยัดนั้นกับเฟรมที่มีกิจกรรมสูงในการผลิตวิดีโอทัศน์ให้มีคุณภาพที่ราบเรียบ เห็นได้จากค่าเบี่ยงเบนมาตรฐานของระเบียบวิธีที่นำเสนอจะมีค่าน้อยกว่า

ขนาดของหน้าต่างที่มากจะให้คุณภาพของสัญญาณวิดีโอทัศน์ที่ราบเรียบกว่าขนาดของหน้าต่างที่มีขนาดเล็กกว่า เฟรมที่มีคุณภาพเพิ่มมากที่สุดของระเบียบวิธีที่นำเสนอที่ขนาดหน้าต่าง 100 จะมีค่ามากกว่าที่ขนาดหน้าต่าง 50 นั้นเป็นเพราะจะมีข้อมูลของเฟรมในอนาคตที่มากกว่านำมาใช้ในการจัดสรรบิตซ้ำ ทั้งนี้ค่าเบี่ยงเบนมาตรฐานก็จะมีค่าน้อยกว่าด้วย ซึ่งจะบ่งบอกถึงว่าเมื่อขนาดหน้าต่างมีขนาดใหญ่กว่า คุณภาพของสัญญาณวิดีโอทัศน์จะเกาะกลุ่มกันมากกว่า เพราะจำนวนบิตจะถูกเฉลี่ยไปในช่วงที่กว้างกว่า และที่ขนาดหน้าต่างทั้ง 50 และ 100 ค่าเบี่ยงเบนมาตรฐานของค่า PSNR จะมีค่าน้อยกว่าวิธีมาตรฐานอีกด้วย ผลของขนาดหน้าต่างต่อขนาดของบัฟเฟอร์ตัวถอดรหัส และเวลาบรจุก่อน ที่ขนาดหน้าต่าง 100 จะต้องการเวลาบรจุก่อน และขนาดของบัฟเฟอร์ตัวถอดรหัสที่มากกว่าที่ขนาดหน้าต่าง 50 เนื่องมาจากจำนวนบิตจะถูกเฉลี่ยไปในช่วงที่กว้างกว่า ทำให้ในหลาย ๆ เฟรมจะมีจำนวนบิตที่มากกว่า ดูได้จากจำนวนอัตราบิตที่ได้จึงทำให้ต้องการเวลาบรจุก่อน และขนาดของบัฟเฟอร์ตัวถอดรหัสที่มากกว่า

ข้อดีของระเบียบวิธีที่นำเสนอสำหรับลำดับวิดีโอทัศน์ Carphone จะมีความต้องการขนาดของบัฟเฟอร์ตัวถอดรหัส และเวลาบรจุก่อนมากกว่าเมื่อเปรียบเทียบกับ JM6.1 เนื่องมาจากเทคนิคที่นำเสนอมีการจัดสรรบิตให้เฉลี่ยไปในหลาย ๆ เฟรม ทำให้ในหลาย ๆ เฟรมจำนวนบิตที่ใช้มีค่ามากกว่าเมื่อเปรียบเทียบกับ JM6.1 ขนาดของบัฟเฟอร์ตัวถอดรหัสที่ต้องการ และเวลาบรจุก่อนจึงมีค่ามากกว่า แต่ดังแสดงในตาราง บัฟเฟอร์ตัวถอดรหัสที่ต้องการโดยเฉลี่ยเพียงไม่กี่ร้อยกิโลบิต และเวลาบรจุก่อนที่ต้องการก็มีค่าน้อยกว่า 1 วินาที ซึ่งเป็นเหตุเป็นผลสำหรับในงานประยุกต์วิดีโอทัศน์ทั่ว ๆ ไป

4.2.2.3 ผลการทดสอบกับลำดับสัญญาณวิดีโอ Suzie+Trevor

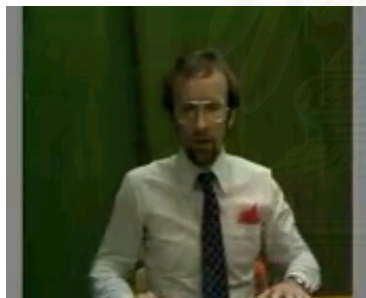
ก) อัตราบิตเป้าหมาย 32 กิโลบิตต่อวินาที และขนาดหน้าต่าง $W=50$ และ 100

ผลการเปรียบเทียบคุณภาพของสัญญาณวิดีโอของลำดับวิดีโอ Suzie+Trevor ที่อัตราบิตเป้าหมาย 32 กิโลบิตต่อวินาที ที่ขนาดหน้าต่าง $W=50$ และ 100 จะแบ่งเป็นผลการเปรียบเทียบคุณภาพเชิงอัตวิสัย จะเป็นผลการเปรียบเทียบระหว่างเฟรมที่มีค่าคุณภาพเพิ่มมากที่สุด และเฟรมที่มีค่าคุณภาพลดลงมากที่สุดของวิธีที่นำเสนอกับการเข้ารหัสมาตรฐาน H.264 เพื่อแสดงให้เห็นว่าในเฟรมที่มีคุณภาพลดลงมากที่สุดเมื่อมองด้วยตาแล้วจะสังเกตเห็นความแตกต่างได้ยาก และในเฟรมที่มีคุณภาพเพิ่มขึ้นมากที่สุดจะสังเกตเห็นความแตกต่างได้อย่างชัดเจน แสดงดังรูปที่ 4.26 และ 4.27 และผลการเปรียบเทียบคุณภาพเชิงประวิสัยด้วยค่า PSNR แสดงดังรูปที่ 4.28

เฟรมที่มีคุณภาพลดลงมากที่สุด

JM6.1

ระเบียบวิธีที่นำเสนอ



(ก)

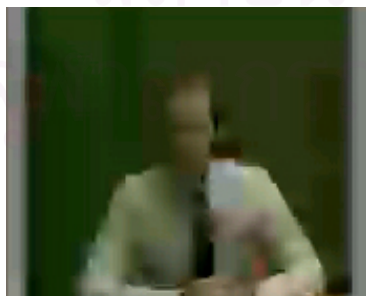


(ข)

เฟรมที่มีคุณภาพเพิ่มมากที่สุด

JM6.1

ระเบียบวิธีที่นำเสนอ



(ค)



(ง)

รูปที่ 4.26 การวัดคุณภาพเชิงอัตวิสัยสำหรับลำดับวิดีโอ Suzie+Trevor ที่อัตราบิตเป้าหมาย 32 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W=50$ เฟรมที่ 87 มีคุณภาพลดลงมากที่สุด = 6.24 dB

สำหรับ JM6.1 แสดงใน (ก) และระเบียบวิธีที่นำเสนอ แสดงใน (ข) เฟรมที่ 70 มีคุณภาพเพิ่มมากขึ้นที่สุด =7.93 dB สำหรับ JM6.1 แสดงใน (ค) และระเบียบวิธีที่นำเสนอ แสดงใน (ง)

เฟรมที่มีคุณภาพลดลงมากที่สุด

JM6.1

ระเบียบวิธีที่นำเสนอ



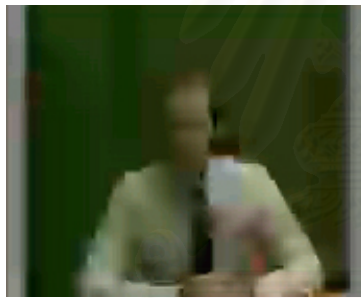
(ก)

(ข)

เฟรมที่มีคุณภาพเพิ่มมากที่สุด

JM6.1

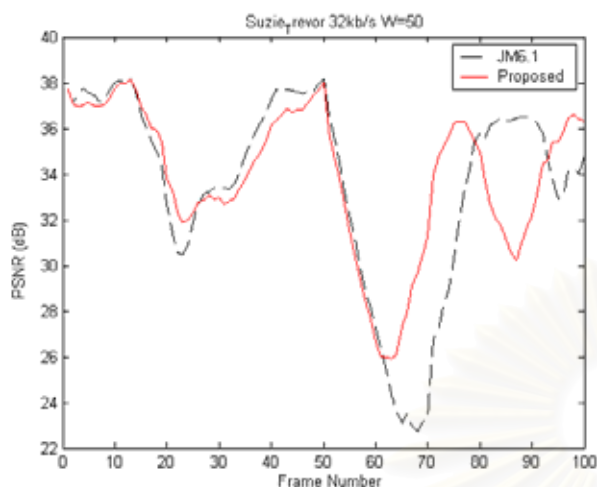
ระเบียบวิธีที่นำเสนอ



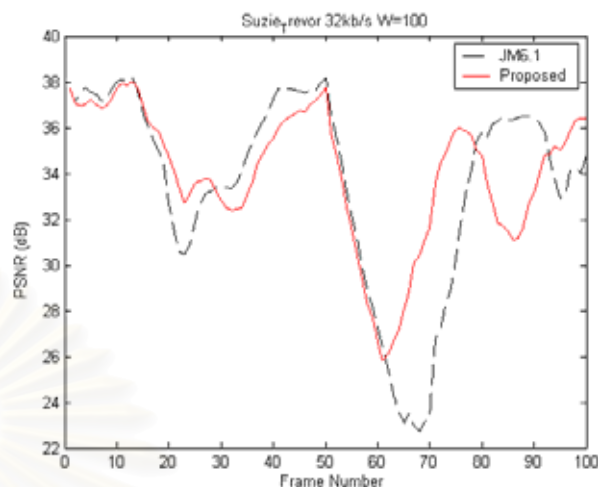
(ค)

(ง)

รูปที่ 4.27 การวัดคุณภาพเชิงอัตวิสัยสำหรับลำดับวิดีโอ Suzie+Trevor ที่อัตราบิตเป้าหมาย 32 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W=100$ เฟรมที่ 87 มีคุณภาพลดลงมากที่สุด =5.26 dB สำหรับ JM6.1 แสดงใน (ก) และระเบียบวิธีที่นำเสนอ แสดงใน (ข) เฟรมที่ 70 มีคุณภาพเพิ่มมากขึ้นที่สุด =8.28 dB สำหรับ JM6.1 แสดงใน (ค) และระเบียบวิธีที่นำเสนอ แสดงใน (ง)



(ก)



(ข)

รูปที่ 4.28 กราฟการเปรียบเทียบคุณภาพลำดับวิดีโอ Suzie+Trevor ที่อัตราบิตเป้าหมาย 32 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W = 50$ แสดงใน (ก) และ ขนาดหน้าต่าง $W = 100$ แสดงใน (ข)

ข) อัตราบิตเป้าหมาย 64 กิโลบิตต่อวินาที และขนาดหน้าต่าง $W = 50$ และ 100

ผลการเปรียบเทียบคุณภาพของสัญญาณวิดีโอของลำดับวิดีโอ Suzie+Trevor ที่อัตราบิตเป้าหมาย 64 กิโลบิตต่อวินาที ที่ขนาดหน้าต่าง $W = 50$ และ 100 จะแบ่งเป็นผลการเปรียบเทียบคุณภาพเชิงอัตวิสัย จะเป็นผลการเปรียบเทียบระหว่างเฟรมที่มีค่าคุณภาพเพิ่มมากที่สุด และเฟรมที่มีค่าคุณภาพลดลงมากที่สุดของวิธีที่นำเสนอกับการเข้ารหัสมาตรฐาน H.264 เพื่อแสดงให้เห็นว่าในเฟรมที่มีคุณภาพลดลงมากที่สุดเมื่อมองด้วยตาแล้วจะสังเกตเห็นความแตกต่างได้ยาก และในเฟรมที่มีคุณภาพเพิ่มขึ้นมากที่สุดจะสังเกตเห็นความแตกต่างได้อย่างชัดเจน แสดงดังรูปที่ 4.29 และ 4.30 และผลการเปรียบเทียบคุณภาพเชิงปริวิสัยด้วยค่า PSNR แสดงดังรูปที่ 4.31

เฟรมที่มีคุณภาพลดลงมากที่สุด

JM6.1



(ก)

ระเบียบวิธีที่นำเสนอ



(ข)

เฟรมที่มีคุณภาพเพิ่มมากที่สุด

JM6.1



(ค)

ระเบียบวิธีที่นำเสนอ



(ง)

รูปที่ 4.29 การวัดคุณภาพเชิงอัตวิสัยสำหรับลำดับวิดีโอ Suzie+Trevor ที่อัตราบิตเป้าหมาย 64 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W=50$ เฟรมที่ 87 มีคุณภาพลดลงมากที่สุด = 1.51 dB สำหรับ JM6.1 แสดงใน (ก) และระเบียบวิธีที่นำเสนอ แสดงใน (ข) เฟรมที่ 64 มีคุณภาพเพิ่มมากขึ้นที่สุด = 5.72 dB สำหรับ JM6.1 แสดงใน (ค) และระเบียบวิธีที่นำเสนอ แสดงใน (ง)

เฟรมที่มีคุณภาพลดลงมากที่สุด

JM6.1



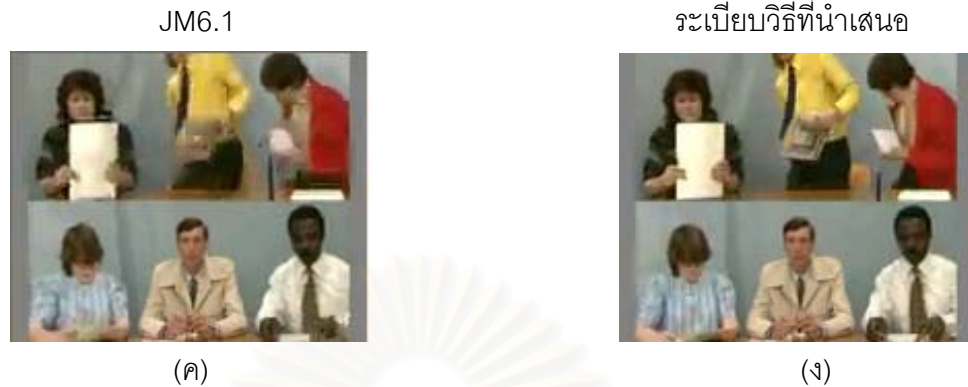
(ก)

ระเบียบวิธีที่นำเสนอ

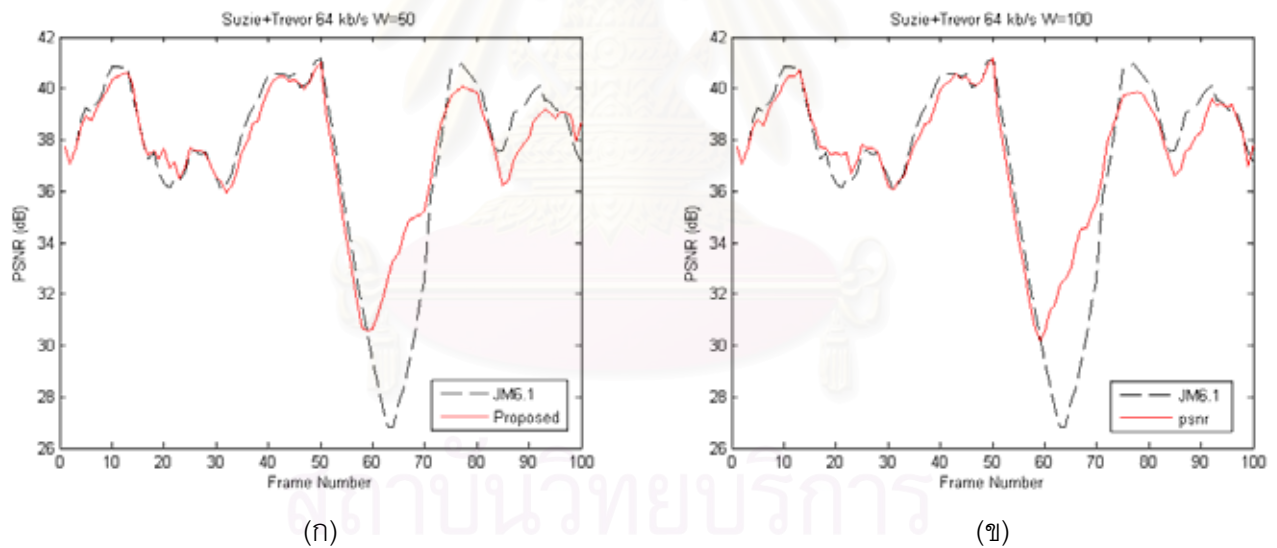


(ข)

เฟรมที่มีคุณภาพเพิ่มมากที่สุด



รูปที่ 4.30 การวัดคุณภาพเชิงอัตตวิสัยสำหรับลำดับวิดีโอ Suzie+Trevor ที่อัตราบิตเป้าหมาย 64 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W=100$ เฟรมที่ 87 มีคุณภาพลดลงมากที่สุด = 1.83 dB สำหรับ JM6.1 แสดงใน (ก) และระเบียบวิธีที่นำเสนอ แสดงใน (ข) เฟรมที่ 64 มีคุณภาพเพิ่มมากขึ้นที่สุด = 6.50 dB สำหรับ JM6.1 แสดงใน (ค) และระเบียบวิธีที่นำเสนอ แสดงใน (ง)



รูปที่ 4.31 กราฟการเปรียบเทียบคุณภาพลำดับวิดีโอ Suzie+Trevor ที่อัตราบิตเป้าหมาย 64 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W=50$ แสดงใน (ก) และ ขนาดหน้าต่าง $W=100$ แสดงใน (ข)

ค) อัตราบิตเป้าหมาย 128 กิโลบิตต่อวินาที และขนาดหน้าต่าง $W=50$ และ 100

ผลการเปรียบเทียบคุณภาพของสัญญาณวิดีโอของลำดับวิดีโอ Suzie+Trevor ที่อัตราบิตเป้าหมาย 128 กิโลบิตต่อวินาที ที่ขนาดหน้าต่าง $W=50$ และ 100 จะแบ่งเป็นผลการเปรียบเทียบคุณภาพเชิงอัตตวิสัย จะเป็นผลการเปรียบเทียบระหว่างเฟรมที่มีค่าคุณภาพเพิ่มมากที่สุด

ที่สุด และเฟรมที่มีค่าคุณภาพลดลงมากที่สุดของวิธีที่นำเสนอกับการเข้ารหัสมาตรฐาน H.264 เพื่อแสดงให้เห็นว่าในเฟรมที่มีคุณภาพลดลงมากที่สุดเมื่อมองด้วยตาแล้วจะสังเกตเห็นความแตกต่างได้ยาก และในเฟรมที่มีคุณภาพเพิ่มขึ้นมากที่สุดจะสังเกตเห็นความแตกต่างได้อย่างชัดเจน แสดงดังรูปที่ 4.32 และ 4.33 และผลการเปรียบเทียบคุณภาพเชิงประจักษ์ด้วยค่า PSNR แสดงดังรูปที่ 4.34

เฟรมที่มีคุณภาพลดลงมากที่สุด

JM6.1



(ก)

ระเบียบวิธีที่นำเสนอ



(ข)

เฟรมที่มีคุณภาพเพิ่มมากที่สุด

JM6.1



(ค)

ระเบียบวิธีที่นำเสนอ



(ง)

รูปที่ 4.32 การวัดคุณภาพเชิงอัตวิสัยสำหรับลำดับวิดีโอ Suzie+Trevor ที่อัตราบิตเป้าหมาย 128 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W=50$ เฟรมที่ 87 มีคุณภาพลดลงมากที่สุด = 2.85 dB สำหรับ JM6.1 แสดงใน (ก) และระเบียบวิธีที่นำเสนอ แสดงใน (ข) เฟรมที่ 59 มีคุณภาพเพิ่มขึ้นที่สุด = 3.58 dB สำหรับ JM6.1 แสดงใน (ค) และระเบียบวิธีที่นำเสนอ แสดงใน (ง)

เฟรมที่มีคุณภาพลดลงมากที่สุด

JM6.1



(ก)

ระเบียบวิธีที่นำเสนอ



(ข)

เฟรมที่มีคุณภาพเพิ่มมากที่สุด

JM6.1



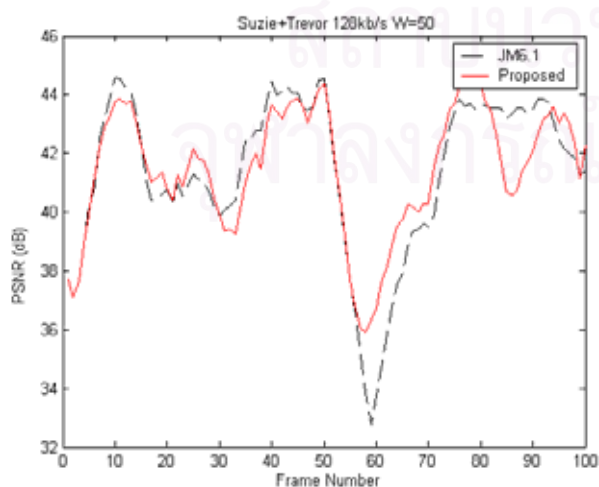
(ค)

ระเบียบวิธีที่นำเสนอ

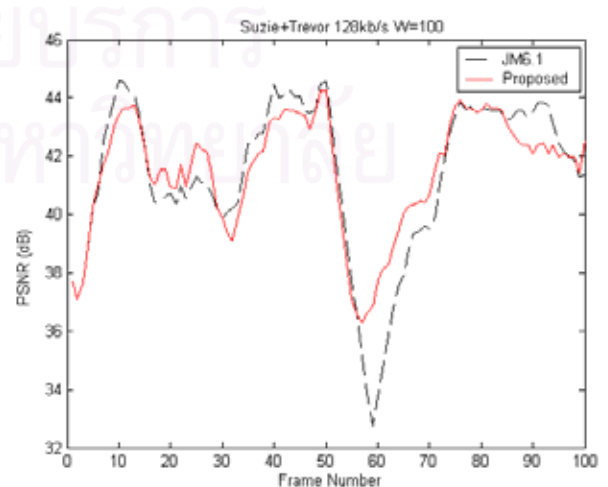


(ง)

รูปที่ 4.33 การวัดคุณภาพเชิงอัตวิสัยสำหรับลำดับวิดีโอ Suzie+Trevor ที่อัตราบิตเป้าหมาย 128 กิโลบิตต่อวินาที ขนาดหน้าต่างต่าง $W=100$ เฟรมที่ 93 มีคุณภาพลดลงมากที่สุด = 1.61 dB สำหรับ JM6.1 แสดงใน (ก) และระเบียบวิธีที่นำเสนอ แสดงใน (ข) เฟรมที่ 59 มีคุณภาพเพิ่มมากขึ้นที่สุด = 4.07 dB สำหรับ JM6.1 แสดงใน (ค) และระเบียบวิธีที่นำเสนอ แสดงใน (ง)



(ก)



(ข)

รูปที่ 4.34 กราฟการเปรียบเทียบคุณภาพลำดับวิดีโอ Suzie+Trevor ที่อัตราบิตเป้าหมาย 128 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W=50$ แสดงใน (ก) และ ขนาดหน้าต่าง $W=100$ แสดงใน (ข)

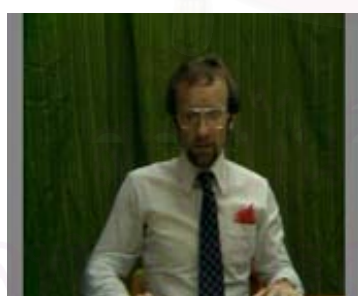
ง) อัตราบิตเป้าหมาย 256 กิโลบิตต่อวินาที และขนาดหน้าต่าง $W=50$ และ 100

ผลการเปรียบเทียบคุณภาพของสัญญาณวิดีโอของลำดับวิดีโอ Suzie+Trevor ที่อัตราบิตเป้าหมาย 256 กิโลบิตต่อวินาที ที่ขนาดหน้าต่าง $W=50$ และ 100 จะแบ่งเป็นผลการเปรียบเทียบคุณภาพเชิงอัตวิสัย จะเป็นผลการเปรียบเทียบระหว่างเฟรมที่มีค่าคุณภาพเพิ่มมากที่สุด และเฟรมที่มีค่าคุณภาพลดลงมากที่สุดของวิธีที่นำเสนอกับการเข้ารหัสมาตรฐาน H.264 เพื่อแสดงให้เห็นว่าในเฟรมที่มีคุณภาพลดลงมากที่สุดเมื่อมองด้วยตาแล้วจะสังเกตเห็นความแตกต่างได้ยาก และในเฟรมที่มีคุณภาพเพิ่มขึ้นมากที่สุดจะสังเกตเห็นความแตกต่างได้อย่างชัดเจน แสดงดังรูปที่ 4.35 และ 4.36 และผลการเปรียบเทียบคุณภาพเชิงปริวิสัยด้วยค่า PSNR แสดงดังรูปที่ 4.37 และตารางที่ 4.4 จะเป็นตารางสรุปผลการเปรียบเทียบค่าคุณภาพ รวมถึงค่าขนาดของบัพเฟอร์ตัวถอดรหัส และค่าเวลาบรจุก่อนที่ต้องการของลำดับวิดีโอ Suzie+Trevor ที่อัตราบิตเป้าหมายทั้งหมดที่ทำการทดสอบ

เฟรมที่มีคุณภาพลดลงมากที่สุด

JM6.1

ระเบียบวิธีที่นำเสนอ



(ก)



(ข)

เฟรมที่มีคุณภาพเพิ่มมากที่สุด

JM6.1



(ค)

ระเบียบวิธีที่นำเสนอ



(ง)

รูปที่ 4.35 การวัดคุณภาพเชิงอัตวิสัยสำหรับลำดับวิดีโอ Suzie+Trevor ที่อัตราบิตเป้าหมาย 256 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W=50$ เฟรมที่ 85 มีคุณภาพลดลงมากที่สุด = 5.82 dB สำหรับ JM6.1 แสดงใน (ก) และระเบียบวิธีที่นำเสนอ แสดงใน (ข) เฟรมที่ 65 มีคุณภาพเพิ่มมากขึ้นที่สุด = 4.43 dB สำหรับ JM6.1 แสดงใน (ค) และระเบียบวิธีที่นำเสนอ แสดงใน (ง)

เฟรมที่มีคุณภาพลดลงมากที่สุด

JM6.1



(ก)

ระเบียบวิธีที่นำเสนอ



(ข)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

เฟรมที่มีคุณภาพเพิ่มมากที่สุด

JM6.1



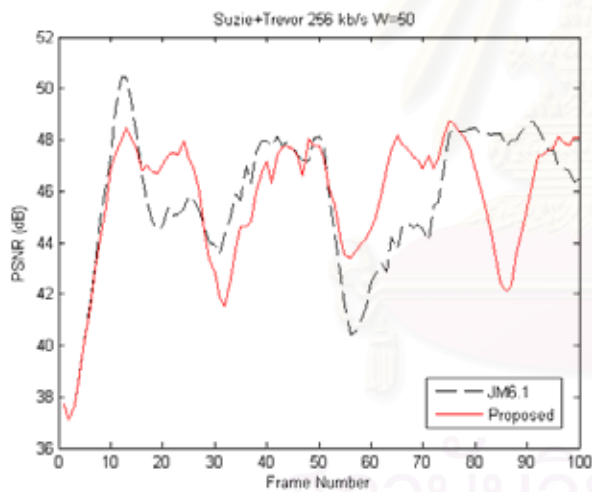
(ค)

ระเบียบวิธีที่นำเสนอ

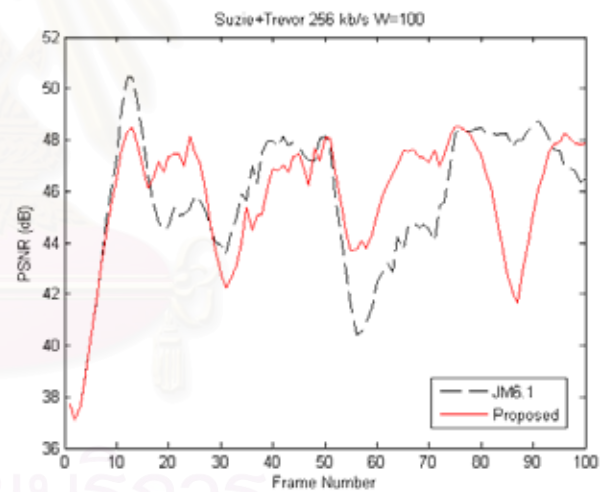


(ง)

รูปที่ 4.36 การวัดคุณภาพเชิงอัตตวิสัยสำหรับลำดับวิดีโอ Suzie+Trevor ที่อัตราบิตเป้าหมาย 256 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W=100$ เฟรมที่ 87 มีคุณภาพลดลงมากที่สุด =6.39 dB สำหรับ JM6.1 แสดงใน (ก) และระเบียบวิธีที่นำเสนอ แสดงใน (ข) เฟรมที่ 65 มีคุณภาพเพิ่มมากขึ้นที่สุด =3.82 dB สำหรับ JM6.1 แสดงใน (ค) และระเบียบวิธีที่นำเสนอ แสดงใน (ง)



(ก)



(ข)

รูปที่ 4.37 กราฟการเปรียบเทียบคุณภาพลำดับวิดีโอ Suzie+Trevor ที่อัตราบิตเป้าหมาย 256 กิโลบิตต่อวินาที ขนาดหน้าต่าง $W=50$ แสดงใน (ก) และ ขนาดหน้าต่าง $W=100$ แสดงใน (ข)

ตารางที่ 4.4 ผลการเปรียบเทียบคุณภาพของลำดับวิดีโอ Suzie+Trevor สำหรับการเข้ารหัส JM6.1 และระเบียบวิธีที่นำเสนอ ที่ขนาดหน้าต่าง $W = 50$ และ 100

Video sequence	Target Bit rate		Average PSNR (dB)	PSNR standard deviation	Average PSNR improvement (dB)	Maximum PSNR improvement (dB)	Maximum PSNR degradation (dB)	Require buffer size (kbits)	Require Pre-loading time (second)	Bit rate (kbps)
Suzie + Trevor	32 kb/s	JM6.1	33.61	4.31	—	—	—	50.85	1.02	32.05
		W=50	33.93	3.63	0.32	7.93	6.24	35.47	0.61	32.34
		W=100	34.01	2.98	0.40	8.28	5.26	37.70	0.66	32.38
	64 kb/s	JM6.1	37.38	3.61	—	—	—	65.52	0.77	64.34
		W=50	37.61	2.52	0.23	5.72	1.51	44.11	0.53	64.97
		W=100	37.61	2.48	0.23	6.50	1.83	44.56	0.55	64.98
	128 kb/s	JM6.1	41.38	2.72	—	—	—	87.12	0.55	128.50
		W=50	41.48	2.15	0.10	3.60	2.85	82.48	0.51	128.32
		W=100	41.50	2.02	0.12	4.10	1.61	96.31	0.60	128.71
	256 kb/s	JM6.1	45.73	2.80	—	—	—	171.80	0.37	257.15
		W=50	45.76	2.56	0.03	4.43	5.82	299.08	0.85	257.33
		W=100	45.79	2.50	0.06	3.82	6.39	304.70	0.86	257.75

จากผลการทดสอบคุณภาพของลำดับวิดีโอ Suzie+Trevor ซึ่งอัดตัวด้วย และเชิงประวัติ เราเห็นว่าระเบียบวิธีที่นำเสนอจะให้คุณภาพที่เพิ่มขึ้นมากกว่าวิธีมาตรฐานในหลาย เฟรม และในเฟรมที่มีคุณภาพลดลงมากที่สุดก็จะสังเกตเห็นได้ยากโดยเฉพาะที่อัตราบิต 32 และ 64 กิโลบิตต่อวินาที ในช่วงเฟรมที่ 50 จนถึงเฟรมที่ 70 ของลำดับวิดีโอ ซึ่งในช่วงที่มีการเปลี่ยนจากลำดับวิดีโอ Suzie เป็นลำดับวิดีโอ Trevor และมีการตัดภาพจากเฟรมที่มีหลายวัตถุเคลื่อนที่ต่างกันออกไป เหลือเพียงวัตถุเดียว ซึ่งมีความซับซ้อนที่สูง ทำให้ค่า PSNR ตกลงอย่างรวดเร็ว เทคนิคที่นำเสนอจะจัดสรรบิตให้มากทำให้คุณภาพที่ได้ดีกว่า สังเกตได้จากเฟรมที่มีคุณภาพเพิ่มขึ้นมากที่สุด ในทั้งเฟรมจะเห็นว่าเทคนิคที่นำเสนอจะให้คุณภาพที่ดีกว่าเมื่อเปรียบเทียบกับ JM6.1 ในเฟรมที่มีค่าความแตกต่างสัมบูรณ์เฉลี่ยสูงสุดจะไม่จำเป็นเสมอไปว่าเทคนิคที่นำเสนอจะให้คุณภาพสูงสุด เพราะจากการทดสอบบิตเป้าหมายจะยังคงมีค่าต่างจากบิตที่ใช้ในการเข้ารหัสจริงค่อนข้างมาก และการคำนวณหาบิตเป้าหมายในระดับเฟรม จะยังคงเป็นบิตเป้าหมายที่มาจากระดับบัพเฟอร์ และบิตเป้าหมายจากเทคนิคที่นำเสนอ แต่จะมีการถ่วงน้ำหนักให้บิตเป้าหมายที่มาจากเทคนิคที่นำเสนอมากกว่าตามสมการ (2.11) บิตเป้าหมายสุดท้ายในระดับเฟรมที่เฟรมที่มีค่าความแตกต่างสัมบูรณ์เฉลี่ยสูงสุด จึงไม่จำเป็นต้องมีค่ามากที่สุด คุณภาพในเฟรมนั้นจึงไม่จำเป็นต้องมีค่ามากที่สุดด้วยเหมือนกัน

จากตารางที่ 4.4 และจากกราฟที่แสดงค่า PSNR จะเห็นว่าระเบียบวิธีที่นำเสนอจะให้คุณภาพโดยเฉลี่ยสูงกว่าวิธีมาตรฐานในช่วง 0.03 ถึง 0.40 dB และเฟรมที่มีคุณภาพเพิ่มขึ้นมากที่สุดเพิ่มขึ้นถึง 8.28 dB ซึ่งเป็นในช่วงที่มีการเปลี่ยนของฉากที่ค่า PSNR มีค่าตกลงอย่างรวดเร็ว ระเบียบวิธีที่นำเสนอจะมีการดึงไม่ให้คุณภาพตกลงไปมากเพราะว่าตัวเข้ารหัสจะมีข้อมูลของเฟรมในอนาคต การเปลี่ยนของฉากหลังซึ่งมีผลให้เกิดความไม่ต่อเนื่องของความซับซ้อนของเนื้อหาวิดีโอทัศน์จะสามารถตรวจพบได้ล่วงหน้า โดยปกติเฟรมที่มีการเปลี่ยนของฉากต้องการบิตจำนวนมาก ในการลดบิตที่ใช้สำหรับเฟรมก่อนหน้า และหลังจากเฟรมที่มีการเปลี่ยนของฉาก ทำให้สามารถเข้ารหัสเฟรมที่มีการเปลี่ยนของฉากได้ด้วยจำนวนบิตที่มากกว่าได้ ผลการทดสอบได้แสดงให้เห็นว่าการลดจำนวนบิตสำหรับเฟรมก่อนหน้า และหลังจากเฟรมที่มีการเปลี่ยนของฉากไม่ทำให้คุณภาพของวิดีโอทัศน์ลดลงอย่างมีนัยสำคัญ และบิตที่ได้จากการประหยัดเหล่านั้นสามารถนำมาใช้ในการปรับปรุงคุณภาพวิดีโอทัศน์ในเฟรมที่มีการเปลี่ยนของฉากได้คุณภาพโดยเฉลี่ยทั้งหมดของลำดับวิดีโอทัศน์เมื่อวัดด้วยค่า PSNR จะมีค่าสูงกว่า เมื่อเปรียบเทียบกับ JM6.1 ถึงแม้ว่าที่อัตราบิตสูง คือ 256 กิโลบิตต่อวินาที ในเฟรมที่มีคุณภาพเพิ่มขึ้นมากที่สุดจะเพิ่มขึ้นน้อยกว่าการลดลงของคุณภาพของเฟรมที่มีคุณภาพลดลงมากที่สุด แต่คุณภาพโดยเฉลี่ยทั้งหมดของลำดับวิดีโอทัศน์เมื่อวัดด้วยค่า PSNR ยังคงมีค่าสูงกว่า เมื่อเปรียบเทียบกับ JM6.1 การจัดสรรบิตซ้ำจะทำการประหยัดบิตจากเฟรมที่มีกิจกรรมต่ำ และใช้บิตที่ได้จากการประหยัดนั้นกับเฟรมที่มีกิจกรรมสูงในการผลิตวิดีโอทัศน์ให้มีคุณภาพที่ราบเรียบ เห็นได้จากค่าเบี่ยงเบนมาตรฐานของระเบียบวิธีที่นำเสนอจะมีค่าน้อยกว่า

ขนาดของหน้าต่างที่มากจะให้คุณภาพของสัญญาณวิดีโอทัศน์ที่ราบเรียบกว่าขนาดของหน้าต่างที่มีขนาดเล็กกว่า เฟรมที่มีคุณภาพเพิ่มมากที่สุดของระเบียบวิธีที่นำเสนอที่ขนาดหน้าต่างต่าง 100 จะมีค่ามากกว่าที่ขนาดหน้าต่าง 50 นั้นเป็นเพราะจะมีข้อมูลของเฟรมในอนาคตที่มากกว่านำมาใช้ในการจัดสรรบิตซ้ำ ทั้งนี้ค่าเบี่ยงเบนมาตรฐานก็จะมีค่าน้อยกว่าด้วย ซึ่งจะบ่งบอกถึงว่าเมื่อขนาดหน้าต่างมีขนาดใหญ่กว่า คุณภาพของสัญญาณวิดีโอทัศน์จะเกาะกลุ่มกันมากกว่า เพราะจำนวนบิตจะถูกเฉลี่ยไปในช่วงที่กว้างกว่า และที่ขนาดหน้าต่างทั้ง 50 และ 100 ค่าเบี่ยงเบนมาตรฐานของค่า PSNR จะมีค่าน้อยกว่าวิธีมาตรฐานอีกด้วย ผลของขนาดหน้าต่างต่อขนาดของบัพเฟอร์ตัวถอดรหัส และเวลาบรจุก่อน ที่ขนาดหน้าต่าง 100 จะต้องการเวลาบรจุก่อน และขนาดของบัพเฟอร์ตัวถอดรหัสที่มากกว่าที่ขนาดหน้าต่าง 50 เนื่องมาจากจำนวนบิตจะถูกเฉลี่ยไปในช่วงที่กว้างกว่า ทำให้ในหลาย ๆ เฟรมจะมีจำนวนบิตที่มากกว่า ดูได้จากจำนวนอัตราบิตที่ได้จึงทำให้ต้องการเวลาบรจุก่อน และขนาดของบัพเฟอร์ตัวถอดรหัสที่มากกว่า

ข้อดีของระเบียบวิธีที่นำเสนอสำหรับลำดับวิดีโอทัศน์ Suzie+Trevor จะมีความต้องการขนาดของบัฟเฟอร์ตัวถอดรหัส และเวลาบรจุก่อนมากกว่าสำหรับที่อัตราบิต 128 และ 256 กิโลบิตต่อวินาที เมื่อเปรียบเทียบกับ JM6.1 เนื่องจากเทคนิคที่นำเสนอมีการจัดสรรบิตให้เฉลี่ยไปในหลาย ๆ เฟรม ทำให้ในหลาย ๆ เฟรมจำนวนบิตที่ใช้มีค่ามากกว่าเมื่อเปรียบเทียบกับ JM6.1 ขนาดของบัฟเฟอร์ตัวถอดรหัสที่ต้องการ และเวลาบรจุก่อนจึงมีค่ามากกว่า แต่ดังแสดงในตาราง บัฟเฟอร์ตัวถอดรหัสที่ต้องการโดยเฉลี่ยเพียงไม่กี่ร้อยกิโลบิต และเวลาบรจุก่อนที่ต้องการก็มีค่าน้อยกว่า 1 วินาที ซึ่งเป็นเหตุเป็นผลสำหรับในงานประยุกต์วิดีโอทัศน์ทั่ว ๆ ไป

4.2.3 ผลการทดสอบเมื่อมีข้อกำหนดของขนาดบัฟเฟอร์ตัวถอดรหัส และเวลาบรจุก่อน

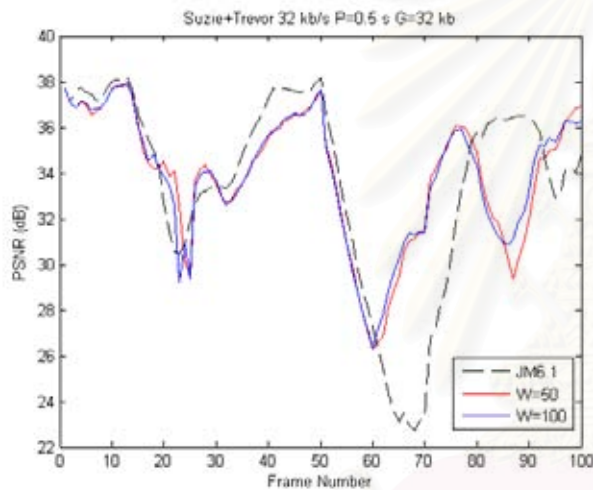
ในหัวข้อนี้เป็นการเปรียบเทียบคุณภาพของสัญญาณวิดีโอทัศน์ของวิธีที่นำเสนอเมื่อมีข้อกำหนดของขนาดบัฟเฟอร์ตัวถอดรหัส และเวลาบรจุก่อน ที่อัตราบิตเป้าหมายต่างๆกัน คือที่ 32 64 128 และ 256 กิโลบิตต่อวินาที และที่ขนาดหน้าต่างที่ต่างกัน คือที่ $W = 50$ และ 100 กับ การเข้ารหัสสัญญาณวิดีโอทัศน์มาตรฐาน H.264 ของลำดับสัญญาณวิดีโอทัศน์ Suzie+Trevor

ในขั้นตอนการทดสอบก่อนหน้า (เมื่อไม่มีข้อกำหนดของขนาดบัฟเฟอร์ตัวถอดรหัส และเวลาบรจุก่อน) ลำดับสัญญาณวิดีโอทัศน์ Suzie+Trevor ที่เข้ารหัสโดยใช้ระเบียบวิธีที่นำเสนอที่ขนาดหน้าต่าง $W = 50$ มีความต้องการเวลาบรจุก่อนในหน่วยวินาที และขนาดของบัฟเฟอร์ตัวถอดรหัสในหน่วยกิโลบิตเท่ากับ 0.61 วินาที 35.47 กิโลบิต สำหรับอัตราบิตเป้าหมาย 32 กิโลบิตต่อวินาที 0.53 วินาที 44.11 กิโลบิต สำหรับอัตราบิตเป้าหมาย 64 กิโลบิตต่อวินาที 0.51 วินาที 82.48 กิโลบิต สำหรับอัตราบิตเป้าหมาย 128 กิโลบิตต่อวินาที และ 0.85 วินาที 299.08 กิโลบิต สำหรับอัตราบิตเป้าหมาย 256 กิโลบิตต่อวินาที และที่ขนาดหน้าต่าง $W = 100$ มีความต้องการเวลาบรจุก่อนในหน่วยวินาที และขนาดของบัฟเฟอร์ตัวถอดรหัสในหน่วยกิโลบิตเท่ากับ 0.66 วินาที 35.47 กิโลบิต สำหรับอัตราบิตเป้าหมาย 32 กิโลบิตต่อวินาที 0.55 วินาที 44.56 กิโลบิต สำหรับอัตราบิตเป้าหมาย 64 กิโลบิตต่อวินาที 0.60 วินาที 96.31 กิโลบิต สำหรับอัตราบิตเป้าหมาย 128 กิโลบิตต่อวินาที และ 0.86 วินาที 304.70 กิโลบิต สำหรับอัตราบิตเป้าหมาย 256 กิโลบิตต่อวินาที

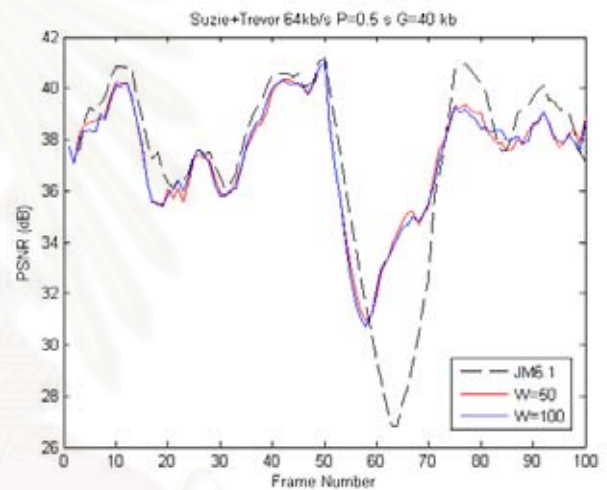
การทดลองในขั้นตอนนี้ จะทำการตั้งค่าข้อกำหนดของขนาดบัฟเฟอร์ที่ตัวถอดรหัสให้มีค่าเท่ากับ 32 40 80 และ 250 กิโลบิต สำหรับอัตราบิตเป้าหมาย 32 64 128 และ 256 กิโลบิตต่อ

วินาที ตามลำดับ และข้อกำหนดของเวลาบรจุก่อนมีค่าเท่ากับ 0.5 วินาที โดยจะใช้ระเบียบวิธีที่นำเสนอในหัวข้อที่ 3.3.1.3 ในการผลิตบิตสตรีมวิดีโอให้เป็นไปตามข้อกำหนดที่ตั้งไว้ (ยกตัวอย่างเช่น ที่อัตราบิตเป้าหมาย 32 กิโลบิตต่อวินาที ผลิตบิตสตรีมวิดีโอที่มีความต้องการขนาดบัพเฟอร์น้อยกว่า 32 กิโลบิต และมีความต้องการเวลาบรจุก่อนน้อยกว่า 0.5 วินาที)

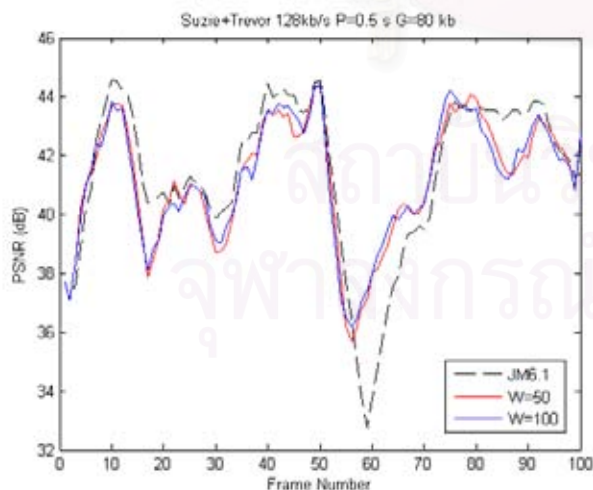
ผลการเปรียบเทียบคุณภาพของสัญญาณวิดีโอของลำดับวิดีโอ Suzie+Trevor จะเป็นผลการเปรียบเทียบคุณภาพเชิงปริวิสัยด้วยค่า PSNR แสดงดังรูปที่ 4.38 และตารางที่ 4.5 จะเป็นตารางสรุปผลการเปรียบเทียบค่าคุณภาพ รวมถึงขนาดสูงสุดของบัพเฟอร์ตัวถอดรหัสที่ได้ และเวลาบรจุก่อนที่ได้ เมื่อมีข้อกำหนดของขนาดบัพเฟอร์ตัวถอดรหัส และเวลาบรจุก่อน



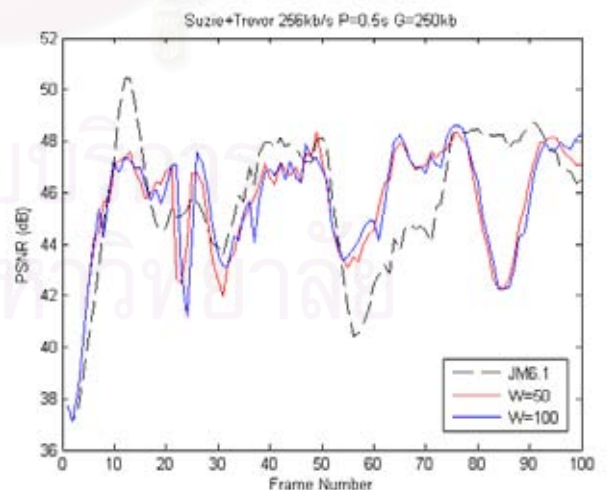
(ก)



(ข)



(ค)



(ง)

รูปที่ 4.38 กราฟการเปรียบเทียบคุณภาพลำดับวิดีโอ Suzie+Trevor ที่ขนาดหน้าต่าง W=50 และ 100 ที่อัตราบิตเป้าหมาย (ก) 32 กิโลบิตต่อวินาที (ข) 64 กิโลบิตต่อวินาที และ (ค) 128

กิโลบิตต่อวินาที (ง) 256 กิโลบิตต่อวินาที เมื่อมีข้อกำหนดของเวลาบรรจุก่อน 0.5 วินาที และมีขนาดของบัฟเฟอร์ตัวถอดรหัส 32 40 80 และ 250 กิโลบิต สำหรับอัตราบิตเป้าหมาย 32 64 128 และ 256 กิโลบิตต่อวินาที ตามลำดับ

ตารางที่ 4.5 ผลการเปรียบเทียบคุณภาพของลำดับวีดิทัศน์ Suzie+Trevor สำหรับการเข้ารหัส JM6.1 และระเบียบวิธีที่นำเสนอที่ขนาดหน้าต่าง $W = 50$ และ 100 เมื่อมีข้อกำหนดของเวลาบรรจุก่อน 0.5 วินาที และมีขนาดของบัฟเฟอร์ตัวถอดรหัส 32 40 80 และ 250 กิโลบิต สำหรับอัตราบิตเป้าหมาย 32 64 128 และ 256 กิโลบิตต่อวินาที ตามลำดับ

Video sequence	Target Bit rate		Average PSNR (dB)	PSNR standard deviation	Average PSNR improvement (dB)	Maximum PSNR improvement (dB)	Maximum PSNR degradation (dB)	Maximum buffer fullness (kbits)	Pre-loading time (second)	Bit rate (kb)
Suzie + Trevor	32 kb/s	JM6.1	33.61	4.31	—	—	—	50.85	1.02	32.05
		W=50	33.87	2.91	0.26	8.43	7.13	33.00	0.54	32.60
		W=100	33.90	2.81	0.29	8.57	5.48	30.34	0.49	32.56
	64 kb/s	JM6.1	37.38	3.61	—	—	—	65.52	0.77	64.34
		W=50	37.34	2.30	-0.04	7.35	2.30	38.20	0.45	64.58
		W=100	37.38	2.30	0.00	7.51	1.94	39.66	0.45	64.75
	128 kb/s	JM6.1	41.38	2.72	—	—	—	87.12	0.55	128.50
		W=50	41.21	2.10	-0.17	4.35	2.53	80.50	0.51	128.38
		W=100	41.20	2.06	-0.18	4.65	2.30	79.50	0.49	128.23
	256 kb/s	JM6.1	45.73	2.80	—	—	—	171.80	0.37	257.15
		W=50	45.56	2.31	-0.17	4.20	5.96	245.49	0.50	257.04
		W=100	45.62	2.31	-0.11	4.50	6.01	248.33	0.49	257.12

จากผลการทดสอบดังแสดงในรูปที่ 4.38 และตารางที่ 4.5 ระเบียบวิธีที่นำเสนอโดยเฉพาะที่อัตราบิตต่ำจะยังคงให้คุณภาพโดยเฉลี่ยสูงกว่าวิธีมาตรฐาน แต่สูงขึ้นน้อยกว่าในขั้นตอนที่แล้ว ทั้งยังเป็นไปตามข้อกำหนดของขนาดของบัฟเฟอร์ที่ตัวถอดรหัส และเวลาบรรจุก่อน ถึงแม้ว่าในอัตราบิตที่สูงขึ้น (64 ถึง 256 กิโลบิตต่อวินาที) จะให้คุณภาพโดยเฉลี่ยต่ำกว่า คือต่ำกว่าในช่วง 0 ถึง 0.18 dB เนื่องมาจากเมื่ออัตราบิตสูงขึ้นระเบียบวิธีที่นำเสนอจะสามารถประหยัดบิตได้มากขึ้นทำให้สามารถกระจายจัดสรรบิตได้ในหลาย ๆ เฟรมได้ดีกว่า ทำให้คุณภาพของวีดิทัศน์ในอัตราบิตสูงมีความราบเรียบกว่า และเนื่องจากมีข้อกำหนดของขนาดบัฟเฟอร์ตัวถอดรหัส และเวลาบรรจุก่อน ทำให้ในหลาย ๆ เฟรมที่มีการกระจายจัดสรรบิตที่ดีกว่าที่อัตราบิตสูง มีโอกาสที่จะไม่เป็นไปตามข้อกำหนดในจำนวนเฟรมที่มากกว่าที่อัตราบิตต่ำ ทำให้ต้องมีกร

ควบคุมให้เป็นไปตามข้อกำหนดในจำนวนเฟรมที่มากกว่าอัตราบิตต่ำ จึงทำให้คุณภาพโดยเฉลี่ยต่ำกว่าวิธีมาตรฐาน แต่ก็ต่ำกว่าเพียงเล็กน้อย โดยจะยังเป็นไปตามข้อกำหนดของขนาดของบัพเฟอร์ และเวลาบรจุก่อน ค่าเบี่ยงเบนมาตรฐานจะยังต่ำกว่า และในช่วงที่มีการเปลี่ยนฉากที่ทำให้ค่า PSNR ลดลงอย่างรวดเร็ว ระเบียบวิธีที่นำเสนอจะยังคงปรับปรุงไม่ให้เกิดลงไปมาก ที่อัตราบิตต่ำ (32 กิโลบิตต่อวินาที) ในเฟรมที่มีคุณภาพเพิ่มขึ้นมากที่สุดจะยังคงสูงกว่าเฟรมที่มีคุณภาพลดลงมากที่สุด คุณภาพสัญญาณวิดีโอที่คืนโดยเฉลี่ยจะยังคงสูงกว่า และยังเป็นไปตามข้อกำหนดของขนาดของบัพเฟอร์ที่ตัวถอดรหัส และเวลาบรจุก่อนที่กำหนดไว้อีกด้วย

บิตสตรีมที่ได้เข้ารหัสจากระเบียบวิธีที่นำเสนอจะยังคงอยู่บนพื้นฐานของมาตรฐาน H.264 และสามารถถอดรหัสได้ด้วยตัวถอดรหัสมาตรฐาน เพื่อหลีกเลี่ยงการล้นของข้อมูลในบัพเฟอร์ตัวถอดรหัส ในมาตรฐาน H.264 จึงมีการกำหนด Hypothetical Reference Decoder (HRD) ขึ้น และบิตสตรีมที่ผลิตจากตัวเข้ารหัสต้องเป็นไปตามข้อกำหนดของ HRD อย่างไรก็ตาม HRD ถูกกำหนดขึ้นสำหรับงานประยุกต์ในเวลาจริง ในงานประยุกต์สตรีมมิงทั่ว ๆ ไป โดยปกติจะมีบัพเฟอร์ตัวถอดรหัสที่ใหญ่มากเพื่อทำให้ค่าประวิงเวลาจิตเตอร์ (delay jitter) ของเครือข่ายมีความราบเรียบ ดังนั้นในงานประเภทนี้ไม่จำเป็นต้องเป็นไปตามข้อกำหนดของ HRD

บทที่ 5

สรุปผลการวิจัย และข้อเสนอแนะ

5.1 สรุปผลการวิจัย

วิทยานิพนธ์นี้นำเสนอเทคนิคการควบคุมอัตราสำหรับการเข้ารหัสวิดีโอที่สนับสนุนระบบออฟไลน์บนพื้นฐานของมาตรฐาน H.264 เพื่อให้คุณภาพของสัญญาณวิดีโอมีค่าดีขึ้น เทคนิคที่นำเสนอนี้จะประยุกต์ใช้ระเบียบวิธีการเข้ารหัสแบบหน้าต่างเลื่อน ซึ่งจะใช้ข้อมูลทางสถิติของเฟรมในอนาคตมาช่วยในการจัดสรรบิต และวิธีของการประมาณบิตเป้าหมายแบบปรับตัวได้ตามความซับซ้อนของการเข้ารหัสในแต่ละเฟรม โดยการใช้อัตราส่วนค่าความแตกต่างสัมบูรณ์เฉลี่ยคุณภาพของลำดับวิดีโอที่ได้อินเฟรมที่มีการเคลื่อนที่สูง และในเฟรมที่มีการเปลี่ยนของฉาก จะดีกว่าลำดับวิดีโอที่เข้ารหัสด้วยมาตรฐาน H.264 JM6.1 การวัดคุณภาพเชิงอัตวิสัย ยังแสดงให้เห็นว่าการปรับปรุงคุณภาพโดยรวมทั้งหมดสามารถทำได้โดยการใช้เทคนิคที่นำเสนอ

ขนาดของบัฟเฟอร์ตัวถอดรหัส และเวลาบรจุก่อนจะถูกนำมาร่วมพิจารณาด้วย เนื่องจากเป็นปัจจัยสำคัญในงานด้านประยุกต์วิดีโอที่สนับสนุน โดยคุณภาพของวิดีโอโดยเฉลี่ยที่ได้จะยังคงดีกว่าลำดับวิดีโอที่เข้ารหัสด้วยมาตรฐาน H.264 JM6.1 และยังเป็นไปตามข้อกำหนดของขนาดของบัฟเฟอร์ตัวถอดรหัส และเวลาบรจุก่อนที่ได้กำหนดมาจากผู้ใช้ ยกเว้นที่อัตราบิตสูง (64 ถึง 256 กิโลบิตต่อวินาที) ที่ให้คุณภาพของวิดีโอโดยเฉลี่ยต่ำกว่า แต่ในช่วงที่มีการเปลี่ยนของฉากทำให้คุณภาพของวิดีโอที่ตกลงอย่างรวดเร็ว เทคนิคที่นำเสนอจะยังสามารถปรับปรุงไม่ให้คุณภาพตกลงไปมาก คุณภาพของวิดีโอที่ได้จะยังคงราบเรียบมากกว่าและยังคงเป็นไปตามข้อกำหนดของขนาดของบัฟเฟอร์ตัวถอดรหัส และเวลาบรจุก่อนที่ได้กำหนดมาจากผู้ใช้อีกด้วย

5.2 ข้อเสนอแนะสำหรับงานวิจัยในอนาคต

งานที่ควรได้รับการศึกษา หรือพัฒนาต่อไปในอนาคต คือ

1. เพิ่มการจัดสรรบิตตามความซับซ้อนของลำดับวิดีโอในระดับหน่วยพื้นฐานย่อย โดยพัฒนาระเบียบวิธีการเข้ารหัสแบบหน้าต่างเลื่อน และวิธีของการประมาณบิต

เป้าหมายแบบปรับตัวได้ ให้สามารถนำไปใช้ในหน่วยพื้นฐานย่อยได้ เพื่อให้การ
จัดสรรบิต และการควบคุมอัตรามีประสิทธิภาพมากขึ้น

2. เพิ่มการเก็บข้อมูลทางสถิติอื่น ๆ ที่ใช้ในการเข้ารหัสมาตรฐาน H.264 ที่สามารถเป็น
ตัวบ่งชี้ความซับซ้อนของลำดับวิดีโอ เช่น โหมดในการเข้ารหัส ขนาดของบล็อกใน
การเข้ารหัส และนำมาประยุกต์ใช้กับระเบียบวิธีการเข้ารหัสแบบหน้าต่างเลื่อน
เพื่อให้จัดสรรบิตตามความซับซ้อนของลำดับวิดีโอได้ดียิ่งขึ้น มีประสิทธิภาพมาก
ขึ้น



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

รายการอ้างอิง

1. Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s. ISO/IEC 11172. (August 1993).
2. Generic coding of moving pictures and associated audio information. ISO/IEC 13818.
3. Video coding for low bit rate communication. ITU-T Recommendation H.263. (March 1996).
4. G. Cote, B. Erol, M. Gallant, and F. Kossentini. H.263+: Video coding at low bit rates. IEEE Transactions on Circuit and Systems For Video Technology. 8 (November 1998) : 849–866.
5. Test model 5 JTC1/SC29/WG11 coding of moving pictures and associated audio MPEG 96/1260. ISO-IEC/JTC1/SC29/WG11. (March 1996).
6. Video codec test model TMN8. ITU-T/SG15. (January 1997).
7. A. R. Reibman and B. G. Haskell. Constraints on variable bit-rate video for ATM networks. IEEE Transactions on Circuit and Systems For Video Technology. 2 (December 1992): 361–372.
8. A. Ortega, K. Ramchandran, and M. Vetterli. Optimal trellis-based buffered compression and fast approximations. IEEE Transactions on Image Processing 3 (January 1994): 26–40.
9. M. Crouse and K. Ramchandran. Joint thresholding and quantizer selection for transform image coding: Entropy-constrained analysis and applications to baseline JPEG. IEEE Transactions on Image Processing. 6, 2 (February 1997): 285–297.
10. M. C. Chen and A. N. Willson Jr. Rate-distortion optimal motion estimation algorithm for motion-compensated transform video coding. IEEE Transactions on Circuit and Systems For Video Technology. 8 (April 1998): 147–158.
11. G. M. Schuster and A. K. Katsaggelos. A theory for the optimal bit allocation between displacement vector field and displaced frame difference. IEEE journal on Selected Areas in Communication. 15 (December 1997): 1739–1751.
12. T. Wiegand, M. Lightstone, D. Mukherjee, T. Campbell, and S. Mitra. Rate-distortion optimized mode selection for very low bit rate video coding and the emerging

- H.263 standard. IEEE Transactions on Circuit and Systems For Video Technology. 6 (April 1996): 182–190.
13. C.-Y. Hsu, A. Ortega, and M. Khansari. Rate control for robust video transmission over burst-error wireless channels. IEEE journal on Selected Areas in Communication. 17 (May 1999): 756–773.
 14. A. Ortega and K. Ramchandran. Rate-distortion methods for image and video compression. IEEE Signal Processing Magazine. 15 (November 1998): 23–50.
 15. W. Ding and B. Liu. Rate control of MPEG video coding and recording by rate-quantization modeling. IEEE Transactions on Circuit and Systems For Video Technology. 6 (February 1996): 12–20.
 16. T. Chiang and Y.-Q. Zhang. A new rate control scheme using quadratic rate distortion model. IEEE Transactions on Circuit and Systems For Video Technology. 7 (February 1997): 246–250.
 17. H.-M. Hang and J.-J. Chen. Source model for transform video coder and its application—Part I: Fundamental theory. IEEE Transactions on Circuit and Systems For Video Technology. 7 (April 1997): 287–298.
 18. J. Ribas-Corbera and S. Lei. Rate control in DCT video coding for low-delay video communications. IEEE Transactions on Circuit and Systems For Video Technology. 9 (February 1999): 172–185.
 19. A. Vetro, H. Sun, and Y. Wang. MPEG-4 rate control for multiple video objects. IEEE Transactions on Circuit and Systems For Video Technology. 9 (February 1999): 186–199.
 20. T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra. Overview of the H.264/AVC Video Coding Standard. IEEE Transactions on Circuit and Systems For Video Technology. 13 (July 2003): 560-576.
 21. พณชิต กิติปัญญางาม. อัลกอริทึมการเข้ารหัสและการค้นหาแบบไฮบริดชนิดปรับตัวได้แบบใช้การปรับย้ายจุดศูนย์กลางสำหรับการประมาณการเคลื่อนที่ในการเข้ารหัสสัญญาณวิดีโอ. วิทยานิพนธ์ปริญญาโท สาขาวิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย 2543.

22. Iain E.G. Richardson. H.264 and MPEG-4 Video Compression. West Sussex. England : Wiley, 2003.
23. D. Marpe, H. Schwarz, and T. Wiegand. Context-Based Adaptive Binary Arithmetic Coding in the H.264/AVC Video Compression Standard. IEEE Transactions on Circuit and Systems For Video Technology. 13 (July 2003): 620-635.
24. JVT-G012r1. Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG Document (March 2003).
25. Z. G. Li, C. Zhu, N. Ling, X. K. Yang, G. N. Feng, S. Wu, and F. Pan. A unified architecture for real-time video-coding systems. IEEE Transactions on Circuit and Systems For Video Technology. 13 (June 2003): 472-487.
26. H.J. Lee, T.H. Chiang and Y.Q. Zhang. Scalable Rate Control for MPEG-4 Video. IEEE Transactions on Circuit and Systems For Video Technology. 10 (September 2000): 878-894.
27. T. Wiegand and B. Girod. Parameter Selection in Lagrangian Hybrid Video Coder Control. Proceeding of International Conference on Image Processing. 3 (October 2001): 542-545.
28. Rao, K. R., and Hwang, J. J. Techniques and Standards for Image, Video and Audio Coding. USA : Prentice Hall, 1996.
29. I.-M. Pao and M.-T. Sun. Encoding Stored Video for Streaming Applications. IEEE Transactions on Circuit and Systems For Video Technology. 11 (February 2001): 199-209.
30. M. Jiang, X. Yi, and N. Ling. Improved Frame Layer Rate Control for H.264 Using MAD Ratio. Proceedings of the 2004 International Symposium on Circuits and Systems. 3 (May 2004): III - 813 -16.
31. ftp://ftp.imtc-files.org/jvt-experts/2003_03_Pattaya/JVT-G012r1_software.zip
32. J. Cai, Z. He, C. W. Chen. Optimal bit allocation for low bit rate video streaming application. Proceeding of International Conference on Image Processing (September 2002): I-73-I-76.
33. Yue Yu, Jian Zhou, Yiliang Wang, Chang Wen Chen. A Novel Two-Pass VBR Coding Algorithm for Fixed-Size Storage Application. IEEE Transactions on Circuit and

Systems For Video Technology. (March 2001): 345 – 356.

34. Harry G. Haskell, Athul Puri, and Arun N. Netravali. Digital Video: an Introduction to MPEG-2. USA: Kluwer Academic Publishersm, 1996.



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ประวัติผู้เขียนวิทยานิพนธ์

นายสิริโรจน์ ก้อนในเมือง เกิดวันที่ 31 ตุลาคม พ.ศ. 2521 ที่จังหวัดนครราชสีมา เข้าศึกษาในหลักสูตรวิศวกรรมศาสตรบัณฑิต คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีสุรนารี ในปีการศึกษา 2539 และเข้าศึกษาต่อในหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต ที่ห้องปฏิบัติการวิจัยกรรมวิธีสัญญาณดิจิทัล ภาควิชาวิศวกรรมไฟฟ้า จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2545



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย