

การควบคุมอัตราการประวิงเวลาต่ำสำหรับมาตรฐานการลงรหัสสัญญาณดิจิทัลขั้นสูง



นายพรเทพ นวกิจกนก

สถาบันวิทยบริการ

จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมไฟฟ้า ภาควิชาวิศวกรรมไฟฟ้า

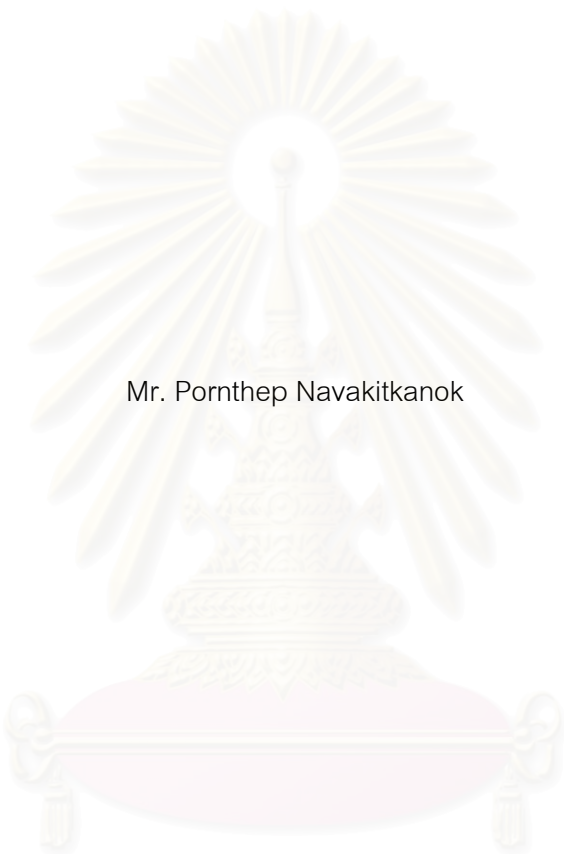
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2546

ISBN 974-17-4338-6

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

LOW DELAY RATE CONTROL FOR ADVANCED VIDEO CODING STANDARD



Mr. Pornthep Navakitkanok

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 2003

ISBN 974-17-4338-6

หัวข้อวิทยานิพนธ์	การควบคุมอัตราการประจุเวลาต่ำสำหรับมาตรฐานการลงรหัส สัญญาณวีดิทัศน์ขั้นสูง
โดย	นายพรเทพ นวกิจกนก
สาขา	วิศวกรรมไฟฟ้า
อาจารย์ที่ปรึกษา	อาจารย์ ดร.สุภาวดี อร่ามวิทย์

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้บัณฑิตวิทยาลัยรับวิทยานิพนธ์ฉบับนี้เป็นส่วน
หนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรบัณฑิต

..... คณบดีคณะวิศวกรรมศาสตร์
(ศาสตราจารย์ ดร.ดิเรก ลาวัณย์ศิริ)

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ
(รองศาสตราจารย์ ดร.สมชาย จิตะพันธ์กุล)

..... อาจารย์ที่ปรึกษา
(อาจารย์ ดร.สุภาวดี อร่ามวิทย์)

..... กรรมการ
(อาจารย์ ดร.สุรีย์ พุ่มรินทร์)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

นายพรเทพ นวกิจกนก : การควบคุมอัตราการประวิงเวลาดำสำหรับมาตรฐานการลงรหัส
 สัญญาณวีดิทัศน์ขั้นสูง. (LOW DELAY RATE CONTROL FOR ADVANCED VIDEO
 CODING STANDARD) อ. ที่ปรึกษา: ดร.สุภาวดี อร่ามวิทย์, 102 หน้า. ISBN 974-17-
 4338-6.

มาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูง มีนัยสำคัญว่า มีประสิทธิภาพการบีบอัด
 สัญญาณวีดิทัศน์ดีกว่ามาตรฐานอื่น ๆ ก่อนหน้าอย่างน้อย 2 เท่า ส่วนประกอบหนึ่งที่มีความสำคัญ
 ในการลงรหัส คือ การควบคุมอัตรา ซึ่งเป็นส่วนที่จำเป็นในการจัดสรรบิตที่เหมาะสมสำหรับแต่ละ
 เฟรม และปรับอัตราบิตที่ผันแปรให้มีอัตราบิตที่ราบเรียบ ในวิทยานิพนธ์นี้ ได้ศึกษาถึงผลกระทบ
 ของการควบคุมอัตราสำหรับการลงรหัสสัญญาณวีดิทัศน์ขั้นสูงเมื่อจำกัดการประวิงเวลาดำ พบว่า
 ค่าคุณภาพสัญญาณวีดิทัศน์ลดลงและมีจำนวนเฟรมกระโดดมาก จึงนำเสนอการปรับปรุงการควบคุม
 อัตราเพื่อเพิ่มประสิทธิภาพในการจัดสรรบิตและการตรวจสอบจำนวนบิตสะสมในบัฟเฟอร์ โดย
 การปรับตัวค่าถ่วงน้ำหนักบิตเป้าหมายและการเงื่อนไขสำหรับการกำหนดค่าควอนไทซ์พารามิเตอร์
 ผลการทดสอบพบว่า วิธีที่นำเสนอไม่มีเฟรมกระโดด และเพิ่มค่าคุณภาพสัญญาณวีดิทัศน์
 ประมาณ 0.3 ถึง 1.19 dB เมื่อเทียบกับการควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวีดิ
 ทัศน์ขั้นสูง

สถาบันวิทยบริการ
 จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชา.....วิศวกรรมไฟฟ้า.....ลายมือชื่อนิสิต.....
 สาขาวิชา.....วิศวกรรมไฟฟ้า.....ลายมือชื่ออาจารย์ที่ปรึกษา.....
 ปีการศึกษา.....2546....

4470710921 : MAJOR ELECTRICAL ENGINEERING

KEY WORD: RATE CONTROL / H.264/MPEG-4 PART 10 / FRAME SKIPPING / LOW DELAY / AVC

PORNTHAP NAVAKITKANOK : LOW DELAY RATE CONTROL FOR ADVANCED VIDEO CODING STANDARD. THESIS ADVISOR: SUPAVADEE ARAMVITH, Ph.D.,102 pp. ISBN 974-17-4338-6.

The upcoming H.264/MPEG-4 Advanced Video Coding (AVC) video compression standard promises a significant improvement in term of providing at least 2x compression improvement over all previous video compression standards. One important component of video codec is rate control. It is a necessary part of an encoder to allocate the suitable number of bits to each video frame and then smooth out the variable bit rate to constant bit-rate channel. In this thesis, we study the impact of low delay constraint on the rate control algorithm proposed for AVC, and show that the performance of AVC rate control is degraded under such low delay constraint. This results in PSNR degradation and many skipped frames. We then propose an improved rate control scheme to effectively allocate the number of bits and monitor the buffer fullness. By adaptively selecting target bit weighting factor and condition for quantization parameter setting, simulation results show that our proposed rate control scheme encodes the video sequences with less frame skipping and with PSNR improvement in the range of 0.3 to 1.19 dB compared to rate control for AVC.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Department Electrical Engineering..... Student's signature.....

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

Academic year ..2003

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยความช่วยเหลืออย่างดียิ่งของท่านอาจารย์ที่ปรึกษาวิทยานิพนธ์ อาจารย์ ดร.สุภาวดี อร่ามวิทย์ ซึ่งได้ให้ความกรุณาแนะนำและเสนอแนะให้ข้อคิดต่างๆในการทำวิจัยมาตลอด ผู้ทำวิจัยจึงขอกราบพระคุณ

ขอขอบคุณหน่วยวิจัยสำหรับศูนย์เชี่ยวชาญทางกรรมวิธีสัญญาณดิจิทัล กองทุนรัชดาสมโภชภิเชก จุฬาลงกรณ์มหาวิทยาลัย ที่ให้ความสนับสนุนในส่วนของคุณและอุปกรณ์ต่างๆ สำหรับช่วยทำงานวิจัย ซึ่งทำให้งานวิจัยดำเนินไปได้ด้วยความราบรื่น

กราบขอบพระคุณบิดา มารดา และครอบครัว ผู้ซึ่งทำให้ผู้วิจัยได้รับแต่สิ่งดีๆ เสมอมา ขอบพระคุณสำหรับความรักความห่วงใยที่ไม่เคยลดน้อยลงเลย

ขอบคุณเพื่อนๆ พี่ๆ น้องๆ ทุกคนที่ผู้วิจัยได้รู้จักไม่ว่าจะเป็นภายในห้องปฏิบัติการ กรรมวิธีสัญญาณดิจิทัลหรือที่ใดก็ตามสำหรับความช่วยเหลือและแรงบันดาลใจที่มีค่าทำให้วิทยานิพนธ์เสร็จสิ้นได้ในที่สุด

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญตาราง.....	ฎ
สารบัญภาพ.....	ฏ
บัญชีคำศัพท์.....	ต
บทที่	
1. บทนำ.....	1
1.1. แนวเหตุผลและความเป็นมา.....	1
1.2. มาตรฐานการลงรหัสสัญญาณวีดิทัศน์.....	3
1.2.1. มาตรฐานของ ITU-T.....	4
1.2.2. มาตรฐานของ ISO/IEC.....	4
1.3. แนวทางที่นำเสนอ.....	6
1.4. วัตถุประสงค์.....	6
1.5. ขอบเขตของวิทยานิพนธ์.....	7
1.6. ขั้นตอนและวิธีการดำเนินงาน.....	7
1.7. ประโยชน์ที่คาดว่าจะได้รับ.....	7
2. ความรู้พื้นฐาน.....	8
2.1. วีดิทัศน์ดิจิทัล (Digital Video).....	8
2.2. การลงรหัสสัญญาณวีดิทัศน์ (Video Coding).....	9
2.2.1. ชนิดของรูปภาพ.....	10
2.2.2. การเข้ารหัสภายในเฟรม.....	11
2.2.3. การเข้ารหัสระหว่างเฟรม.....	15
2.2.4. การแปลงและอินเวอร์สการแปลง(Transform and Inverse Transform).....	16
2.2.5. การควอนไทซ์และการสแกนค่าสัมประสิทธิ์ที่ได้.....	17
2.2.6. การลงรหัสเอนโทรปี.....	17

บทที่	หน้า
2.2.7. การประมาณการเคลื่อนที่ (Motion Estimation) และ การชดเชยการเคลื่อนที่ (Motion Compensation).....	18
2.2.8. รูปแบบการจัดเรียงสัญญาณ.....	22
2.2.9. รูปแบบสัญญาณวีดิทัศน์ที่ใช้.....	23
2.3. การควบคุมอัตราสำหรับการลงรหัสสัญญาณวีดิทัศน์ขั้นสูง.....	23
2.3.1. การควบคุมอัตราระดับกลุ่มภาพ.....	23
2.3.2. การควบคุมอัตราระดับเฟรม.....	26
2.3.2.1. ส่วน Pre-Encoding Stage.....	26
2.3.2.2. ส่วน Post-Encoding Stage.....	29
2.3.3. การควบคุมอัตราระดับหน่วยพื้นฐานย่อย.....	29
2.3.3.1. นิยามของหน่วยพื้นฐานย่อย (Basic unit).....	29
2.3.3.2. การควบคุมอัตราของหน่วยพื้นฐานย่อย.....	30
2.4. การวัดคุณภาพ(Quality Measurement).....	33
2.4.1. ค่าความผิดพลาดกำลังสองเฉลี่ย (Mean Square Error: MSE).....	33
2.4.2. ค่าอัตราส่วนสัญญาณยอดต่อสัญญาณรบกวน (Peak Signal to Noise Ratio: PSNR).....	33
3. รายละเอียดเกี่ยวกับวิทยานิพนธ์และระบบที่นำเสนอ.....	34
3.1. ภาพรวมของระบบ.....	34
3.2. ลักษณะสัญญาณวีดิทัศน์ขาเข้า.....	34
3.3. การทำงานของระบบที่นำเสนอ.....	35
3.3.1. การศึกษาผลกระทบของการสะสมบัฟเฟอร์ของการควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูงต่อคุณภาพสัญญาณวีดิทัศน์เมื่อจำกัดการประวิงเวลา.....	36
3.3.2. ส่วนควบคุมการกระโดดของเฟรม.....	40
3.3.3. ส่วนการควบคุมอัตราในระดับกลุ่มภาพ.....	42
3.3.4. ส่วนการควบคุมอัตราในระดับเฟรม.....	43
3.3.4.1. ส่วน Pre-Encoding Stage.....	43
3.3.4.2. ส่วน Post-Encoding Stage.....	45
3.3.5. การควบคุมอัตราในระดับหน่วยพื้นฐานย่อย.....	46

บทที่	หน้า
4. ผลการทดสอบ.....	48
4.1. ขั้นตอนการทดสอบ.....	48
4.1.1. การกำหนดพารามิเตอร์ในการลงรหัส.....	48
4.1.2. ลำดับภาพที่ใช้ในการทดสอบ.....	49
4.2. ผลการทดสอบและวิเคราะห์ผลการลงรหัส.....	49
4.2.1. ผลการทดสอบเมื่อปรับค่า β ในช่วง 0 ถึง 1 (0, 0.1, 0.2, ..., 1).....	50
4.2.1.1. ผลการทดสอบกับลำดับภาพ Carphone.....	50
4.2.1.2. ผลการทดสอบกับลำดับภาพ Foreman.....	52
4.2.1.3. ผลการทดสอบกับลำดับภาพ Silent.....	53
4.2.1.4. ผลการทดสอบกับลำดับภาพ Salesman.....	54
4.2.2. ผลการทดสอบการกำหนดเงื่อนไขสำหรับการหาค่าควอนไทล์พารามิเตอร์โดย การหาค่าจุดเริ่มเปลี่ยนบัพเฟอร์ที่เหมาะสมเมื่อกำหนดค่าวงน้ำหนักแบบคงที่ และแบบปรับตัวได้.....	55
4.2.2.1. ผลการทดสอบกับลำดับภาพ Carphone.....	55
4.2.2.2. ผลการทดสอบกับลำดับภาพ Foreman.....	58
4.2.2.3. ผลการทดสอบกับลำดับภาพ Silent.....	63
4.2.2.4. ผลการทดสอบกับลำดับภาพ Salesman.....	66
4.2.3. เปรียบเทียบการควบคุมอัตราที่นำเสนอกับการควบคุมอัตราการลงรหัสสัญญาณ วีดิทัศน์ขั้นสูง.....	70
4.2.3.1. ผลการทดสอบกับลำดับภาพ Carphone.....	70
4.2.3.2. ผลการทดสอบกับลำดับภาพ Foreman.....	73
4.2.3.3. ผลการทดสอบกับลำดับภาพ Silent.....	75
4.2.3.4. ผลการทดสอบกับลำดับภาพ Salesman.....	79
5. สรุปผลการวิจัยและข้อเสนอแนะ.....	84
5.1. สรุปผลการวิจัย.....	84
5.2. ข้อเสนอแนะสำหรับงานวิจัยในอนาคต.....	85
รายการอ้างอิง.....	86
ภาคผนวก.....	89

บทที่	หน้า
บทความที่ได้รับการเผยแพร่.....	90
ประวัติผู้เขียนวิทยานิพนธ์.....	102



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญตาราง

ตาราง	หน้า
1.1 มาตรฐานการลงทะเบียนสัญญาณวิทยุของ ITU-T.....	3
1.2 มาตรฐานการลงทะเบียนสัญญาณวิทยุของ ISO/IEC.....	4
2.1 รูปแบบของสัญญาณวิทยุที่ใช้งานในปัจจุบัน.....	8
3.1 ค่าคุณภาพสัญญาณวิทยุและจำนวนเฟรมกระโดดของการควบคุมอัตราสำหรับ มาตรฐานการลงทะเบียนสัญญาณวิทยุขั้นสูงและ การควบคุมอัตราสำหรับมาตรฐานการลงทะเบียนสัญญาณวิทยุขั้นสูงเมื่อจำกัดการประวิงเวลาพารามิเตอร์ที่ใช้ในการทดสอบ.....	41
4.1 พารามิเตอร์ที่ใช้ในการทดสอบ.....	49
4.2 ค่าคุณภาพสัญญาณวิทยุและจำนวนเฟรมกระโดดที่ค่าถ่วงน้ำหนักบิตเป้าหมายต่าง ๆ ของลำดับภาพ Carphone.....	51
4.3 ค่าคุณภาพสัญญาณวิทยุและจำนวนเฟรมกระโดดที่ค่าถ่วงน้ำหนักบิตเป้าหมายต่าง ๆ ของลำดับภาพ Foreman.....	52
4.4 ค่าคุณภาพสัญญาณวิทยุและจำนวนเฟรมกระโดดที่ค่าถ่วงน้ำหนักบิตเป้าหมายต่าง ๆ ของลำดับภาพ Silent.....	53
4.5 ค่าคุณภาพสัญญาณวิทยุและจำนวนเฟรมกระโดดที่ค่าถ่วงน้ำหนักบิตเป้าหมายต่าง ๆ ของลำดับภาพ Salesman.....	54
4.6 จำนวนเฟรมกระโดด จำนวนบิตสะสมในบัพเฟอร์ และค่าคุณภาพสัญญาณวิทยุที่ค่าจุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ ของลำดับภาพ Carphone เมื่อกำหนดค่าถ่วงน้ำหนักเป้าหมายแบบคงที่และแบบปรับตัวได้.....	59
4.7 จำนวนเฟรมกระโดด จำนวนบิตสะสมในบัพเฟอร์ และค่าคุณภาพสัญญาณวิทยุที่ค่าจุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ ของลำดับภาพ Foreman เมื่อกำหนดค่าถ่วงน้ำหนักเป้าหมายแบบคงที่และแบบปรับตัวได้.....	62
4.8 จำนวนเฟรมกระโดด จำนวนบิตสะสมในบัพเฟอร์ และค่าคุณภาพสัญญาณวิทยุที่ค่าจุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ ของลำดับภาพ Silent เมื่อกำหนดค่าถ่วงน้ำหนักเป้าหมายแบบคงที่และแบบปรับตัวได้.....	66
4.9 จำนวนเฟรมกระโดด จำนวนบิตสะสมในบัพเฟอร์ และค่าคุณภาพสัญญาณวิทยุที่ค่าจุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ ของลำดับภาพ Salesman เมื่อกำหนดค่าถ่วงน้ำหนักเป้าหมายแบบคงที่และแบบปรับตัวได้.....	70

4.10 เปรียบเทียบจำนวนเฟรมกระโดดและค่าคุณภาพสัญญาณวีดิทัศน์ระหว่างการควบคุม
อัตราสำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูง (AVC) กับ การควบคุมอัตราที่
นำเสนอ.....83



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญภาพ

ภาพประกอบ	หน้า
รูปที่ 2.1 บล็อกไดอะแกรมของการเข้ารหัสสัญญาณวีดิทัศน์ที่มีการบีบอัด.....	9
รูปที่ 2.2 ความสัมพันธ์ระหว่างเฟรมแบบ I,P และ B ในการทำนายการเคลื่อนที่.....	11
รูปที่ 2.3 ส่วนประกอบความส่องสว่างขนาด 4 x 4 จุดภาพ.....	12
รูปที่ 2.4 ตำแหน่งส่วนประกอบความส่องสว่างขนาด 4 x 4 จุดภาพ.....	12
รูปที่ 2.5 การเข้ารหัสภายในเฟรมของส่วนประกอบความส่องสว่างขนาด 4 x 4 จุดภาพ โหมด 0 ถึง 8	14
รูปที่ 2.6 การเข้ารหัสภายในเฟรมของส่วนประกอบความส่องสว่างขนาด 16 x 16 จุดภาพ โหมด 0 ถึง 3	15
รูปที่ 2.7 การจัดเรียงค่าสัมประสิทธิ์.....	16
รูปที่ 2.8 วิธีการถอดค่าสัมประสิทธิ์แบบซิกแซกและแบบคู่.....	17
รูปที่ 2.9 การแบ่งบล็อกขนาด 16 x 16 จุดภาพ.....	16
รูปที่ 2.10 การแบ่งบล็อกขนาด 8 x 8 จุดภาพ.....	17
รูปที่ 2.11 การแบ่งบล็อกภายในภาพ.....	19
รูปที่ 2.12 การประมาณการเคลื่อนที่	20
รูปที่ 2.13 การหาค่าจุดภาพระหว่าง 2 จุดภาพ.....	20
รูปที่ 2.14 การจัดกลุ่มมาโครบล็อกแบบยืดหยุ่น.....	21
รูปที่ 2.15 การเรียงลำดับมาโครบล็อกแบบปกติ.....	22
รูปที่ 2.16 บล็อกไดอะแกรมการควบคุมอัตราสำหรับการลงรหัสสัญญาณวีดิทัศน์ขั้นสูง.....	24
รูปที่ 2.17 ขนาดหน่วยพื้นฐานย่อย.....	30
รูปที่ 3.1 ภาพรวมระบบ.....	34
รูปที่ 3.2 สัญญาณวีดิทัศน์ที่ใช้ทดสอบในวิทยานิพนธ์.....	35
รูปที่ 3.3 ปัจจัยต่างๆที่มีผลต่อการประวิงเวลาในระบบสื่อสารวีดิทัศน์.....	37
รูปที่ 3.4 ผลการทดสอบการเข้ารหัสสัญญาณวีดิทัศน์ของลำดับภาพ Carphone.....	38
รูปที่ 3.5 ผลการทดสอบการเข้ารหัสสัญญาณวีดิทัศน์ของลำดับภาพ Foreman	39
รูปที่ 3.6 ผลการทดสอบการเข้ารหัสสัญญาณวีดิทัศน์ของลำดับภาพ Silent	39
รูปที่ 3.7 เปรียบเทียบการเข้ารหัสลำดับภาพ Foreman.....	42
รูปที่ 3.8 ค่าคุณภาพสัญญาณวีดิทัศน์และจำนวนเฟรมกระโดดเมื่อปรับค่าถ่วงน้ำหนักบิต เป้าหมายต่าง ๆ.....	44

บทที่	หน้า
รูปที่ 3.9 จำนวนบิตเป้าหมาย และจำนวนบิตที่เข้ารหัสจริงสำหรับลำดับภาพ Carphone.....	46
รูปที่ 3.10 การควบคุมอัตราในระดับหน่วยพื้นฐานย่อยที่น่าเสนอ.....	47
รูปที่ 4.1 ค่าคุณภาพสัญญาณวีดิทัศน์เฉลี่ยที่ค่าถ่วงน้ำหนักบิตเป้าหมายต่าง ๆ ของลำดับภาพ Carphone.....	51
รูปที่ 4.2 ค่าคุณภาพสัญญาณวีดิทัศน์เฉลี่ยที่ค่าถ่วงน้ำหนักบิตเป้าหมายต่าง ๆ ของลำดับภาพ Foreman.....	52
รูปที่ 4.3 ค่าคุณภาพสัญญาณวีดิทัศน์เฉลี่ยที่ค่าถ่วงน้ำหนักบิตเป้าหมายต่าง ๆ ของลำดับภาพ Silent.....	53
รูปที่ 4.4 ค่าคุณภาพสัญญาณวีดิทัศน์เฉลี่ยที่ค่าถ่วงน้ำหนักบิตเป้าหมายต่าง ๆ ของลำดับภาพ Salesman.....	54
รูปที่ 4.5 จำนวนบิตสะสมในบัพเฟอร์ของลำดับภาพ Carphone ที่จุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายคงที่.....	56
รูปที่ 4.6 ค่าคุณภาพสัญญาณวีดิทัศน์ของลำดับภาพ Carphone ที่จุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายคงที่.....	56
รูปที่ 4.7 จำนวนบิตสะสมในบัพเฟอร์ของลำดับภาพ Carphone ที่จุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายปรับตัวได้.....	57
รูปที่ 4.8 ค่าคุณภาพสัญญาณวีดิทัศน์ของลำดับภาพ Carphone ที่จุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายปรับตัวได้.....	57
รูปที่ 4.9 ค่าถ่วงน้ำหนักบิตเป้าหมายของลำดับภาพ Carphone	58
รูปที่ 4.10 จำนวนบิตสะสมในบัพเฟอร์ของลำดับภาพ Foreman ที่จุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายปรับคงที่.....	59
รูปที่ 4.11 ค่าคุณภาพสัญญาณวีดิทัศน์ของลำดับภาพ Foreman ที่จุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายคงที่.....	60
รูปที่ 4.12 จำนวนบิตสะสมในบัพเฟอร์ของลำดับภาพ Foreman ที่จุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายปรับตัวได้.....	60
รูปที่ 4.13 ค่าคุณภาพสัญญาณวีดิทัศน์ของลำดับภาพ Foreman ที่จุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายปรับตัวได้.....	61
รูปที่ 4.14 ค่าถ่วงน้ำหนักบิตเป้าหมายของลำดับภาพ Foreman	62

บทที่	หน้า
รูปที่ 4.15 จำนวนบิตสะสมในบัพเฟอร์ของลำดับภาพ Silent ที่จุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ เมื่อ กำหนดค่าถ่วงน้ำหนักบิตเป้าหมายคงที่.....	63
รูปที่ 4.16 ค่าคุณภาพสัญญาณวีดิทัศน์ของลำดับภาพ Silent ที่จุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ เมื่อ กำหนดค่าถ่วงน้ำหนักบิตเป้าหมายคงที่.....	64
รูปที่ 4.17 จำนวนบิตสะสมในบัพเฟอร์ของลำดับภาพ Silent ที่จุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ เมื่อ กำหนดค่าถ่วงน้ำหนักบิตเป้าหมายปรับตัวได้.....	64
รูปที่ 4.18 ค่าคุณภาพสัญญาณวีดิทัศน์ของลำดับภาพ Silent ที่จุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ เมื่อ กำหนดค่าถ่วงน้ำหนักบิตเป้าหมายปรับตัวได้.....	65
รูปที่ 4.19 ค่าถ่วงน้ำหนักบิตเป้าหมายของลำดับภาพ Silent.....	65
รูปที่ 4.20 จำนวนบิตสะสมในบัพเฟอร์ของลำดับภาพ Salesman ที่จุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆเมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายคงที่.....	67
รูปที่ 4.21 ค่าคุณภาพสัญญาณวีดิทัศน์ของลำดับภาพ Salesman ที่จุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆเมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายคงที่.....	67
รูปที่ 4.22 จำนวนบิตสะสมในบัพเฟอร์ของลำดับภาพ Salesman ที่จุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆเมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายปรับตัวได้.....	68
รูปที่ 4.23 ค่าคุณภาพสัญญาณวีดิทัศน์ของลำดับภาพ Salesman ที่จุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆเมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายปรับตัวได้.....	69
รูปที่ 4.24 ค่าถ่วงน้ำหนักบิตเป้าหมายของลำดับภาพ Salesman.....	69
รูปที่ 4.25 ค่าคุณภาพสัญญาณวีดิทัศน์ของลำดับภาพ Carphone.....	71
รูปที่ 4.26 จำนวนบิตสะสมในบัพเฟอร์ของลำดับ Carphone.....	72
รูปที่ 4.27 จำนวนบิตเป้าหมายและจำนวนบิตที่เข้ารหัสจริงของลำดับภาพ Carphone.....	72
รูปที่ 4.28 เปรียบเทียบการควบคุมอัตราสำหรับการลงรหัสสัญญาณวีดิทัศน์กับการควบคุมอัตราที่ นำเสนอของลำดับภาพ Carphone	74
รูปที่ 4.29 ค่าคุณภาพสัญญาณวีดิทัศน์ของลำดับภาพ Foreman.....	74
รูปที่ 4.30 จำนวนบิตสะสมในบัพเฟอร์ของลำดับ Foreman.....	74
รูปที่ 4.31 จำนวนบิตเป้าหมายและจำนวนบิตที่เข้ารหัสจริงของลำดับภาพ Foreman.....	75
รูปที่ 4.32 เปรียบเทียบการควบคุมอัตราสำหรับการลงรหัสสัญญาณวีดิทัศน์กับการควบคุมอัตราที่ นำเสนอของลำดับภาพ Foreman.....	76
รูปที่ 4.33 ค่าคุณภาพสัญญาณวีดิทัศน์ของลำดับภาพ Silent.....	77

บทที่	หน้า
รูปที่ 4.34 จำนวนบิตสะสมในบัพเฟออร์ของลำดับ Silent.....	77
รูปที่ 4.35 จำนวนบิตเป้าหมายและจำนวนบิตที่เข้ารหัสจริงของลำดับภาพ Silent.....	78
รูปที่ 4.36 เปรียบเทียบการควบคุมอัตราสำหรับการลงรหัสสัญญาณวีดิทัศน์กับการควบคุมอัตราที่ นำเสนอของลำดับภาพ Silent	79
รูปที่ 4.37 ค่าคุณภาพสัญญาณวีดิทัศน์ของลำดับภาพ Salesman.....	80
รูปที่ 4.38 จำนวนบิตสะสมในบัพเฟออร์ของลำดับ Salesman.....	80
รูปที่ 4.39 จำนวนบิตเป้าหมายและจำนวนบิตที่เข้ารหัสจริงของลำดับภาพ Salesman.....	81
รูปที่ 4.40 เปรียบเทียบการควบคุมอัตราสำหรับการลงรหัสสัญญาณวีดิทัศน์กับการควบคุมอัตราที่ นำเสนอของลำดับภาพ Salesman.....	82



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 1

บทนำ

1.1 แนวเหตุผลและความเป็นมา

เนื่องจากข้อมูลด้านสื่อประสม (Multimedia) เช่น ข้อมูลตัวอักษร เสียงพูด ออดิโอ ภาพนิ่ง และวีดิทัศน์มีขนาดใหญ่โดยเฉพาะอย่างยิ่งวีดิทัศน์ดิจิทัล ช่องสัญญาณที่มีการกำหนดให้ในระบบสื่อสารโทรคมนาคมแต่ละระบบมีขนาดจำกัด จะทำให้เกิดความล่าช้าหรือการประวิงเวลา รวมทั้งอาจเกิดการสูญหายของข้อมูล ซึ่งอาจก่อให้เกิดความเสียหายในการสื่อสารได้ เพื่อให้สามารถส่งข้อมูลผ่านช่องสัญญาณได้อย่างมีประสิทธิภาพมากยิ่งขึ้น จึงได้มีการใช้เทคโนโลยีการบีบอัดข้อมูลวีดิทัศน์ดิจิทัล ทำให้สามารถส่งสัญญาณวีดิทัศน์บนช่องสัญญาณที่มีอัตราบิตต่ำอย่างเช่น โครงข่ายโทรศัพท์แบบสวิตซ์สาธารณะ (Public Switched Telephone Network) โครงข่ายพื้นที่ท้องถิ่น (Local Area Network) และโครงข่ายไร้สาย (Wireless Network) ซึ่งสามารถนำไปประยุกต์ใช้กับงานด้านวีดิทัศน์ดิจิทัลได้ อย่างเช่น การประชุมสัมมนาทางวีดิทัศน์ (Videoconference) เป็นต้น

มาตรฐานการบีบอัดข้อมูลสำหรับสัญญาณวีดิทัศน์ที่ใช้กันอย่างแพร่หลาย คือ H.263 [1] MPEG-1 [2] MPEG-2 [3] และ MPEG-4 [4] โดยแต่ละมาตรฐานเหมาะสำหรับการประยุกต์ใช้งานแตกต่างกันไป สำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูง เกิดขึ้นจากการร่วมมือกันระหว่างองค์กร ITU-T (International Telecommunication Union – Telecommunication Sector) และ ISO/IEC (International Standard Organization) โดยออกเป็นมาตรฐานภายใต้องค์กร ITU-T ใช้ชื่อว่า H.264 และ ISO/IEC ออกเป็นมาตรฐานภายใต้ชื่อว่า การลงรหัสสัญญาณวีดิทัศน์ขั้นสูง H.264/MPEG-4 PART 10 (Advanced Video Coding: AVC)

ส่วนประกอบหนึ่งที่สำคัญสำหรับการลงรหัสสัญญาณวีดิทัศน์ คือ การควบคุมอัตรา (Rate control) เนื่องจากสัญญาณวีดิทัศน์ที่ถูกบีบอัด จะมีอัตราบิตผันแปรแต่ละเฟรมขึ้นกับความซับซ้อนของแต่ละเฟรมวีดิทัศน์ที่ไม่เท่ากัน โดยทั่วไปการสื่อสารสัญญาณวีดิทัศน์จะส่งผ่านช่องสัญญาณคงที่ โดยบัพเฟอร์ของตัวเข้ารหัสมีหน้าที่ปรับอัตราบิตให้ราบเรียบเพื่อเหมาะสำหรับช่องสัญญาณที่มีแบนวิidthคงที่ โดยการควบคุมอัตราจะทำการหาค่าควอนไทซ์พารามิเตอร์โดยคำนึงถึงจำนวนบิตสะสมในบัพเฟอร์ แบนด์วิidthของช่องสัญญาณ และความซับซ้อนของเฟรม

การหาค่าควอนไทล์พารามิเตอร์ที่เหมาะสมมีผลต่อค่าคุณภาพสัญญาณวีดิทัศน์โดยตรง ดังนั้นจึงได้มีการศึกษา และวิจัยการควบคุมอัตราจำนวนมาก วิธีการหาค่าควอนไทล์พารามิเตอร์ที่เหมาะสม [14] แบ่งได้เป็น 2 วิธี วิธีโปรแกรมพลวัต (Dynamic programming) เพื่อหาค่าควอนไทล์พารามิเตอร์ที่เหมาะสมโดยทดสอบทุกระดับควอนไทล์ การจำลองความสัมพันธ์ระหว่างอัตราบิตกับความเพี้ยนของข้อมูล (distortion) โดยใช้ตัวคูณลากรองจ์ (Lagrange Multiplier) ถึงแม้ค่าควอนไทล์พารามิเตอร์ที่ได้มาจะไม่ใช่ว่าควอนไทล์พารามิเตอร์ที่เหมาะสมที่สุด แต่วิธีนี้มีความซับซ้อนน้อยกว่าแบบโปรแกรมพลวัต

อีกวิธีหนึ่งใช้โมเดลทางคณิตศาสตร์ [5] [6] โดยการหาค่าควอนไทล์พารามิเตอร์ที่สอดคล้องกับค่าทางสถิติของอินพุท อัตรา และโมเดลความเพี้ยน ในปัจจุบันการควบคุมอัตราด้วยวิธีนี้เป็นที่นิยมเนื่องจากเรียบง่ายไม่ซับซ้อน

โดย [5] เสนอการควบคุมอัตราเมื่อจำกัดการประวิงเวลาสำหรับมาตรฐาน H.263 โดยนำเสนอการควบคุมอัตราในระดับมาโครบล็อกทำให้สามารถการควบคุมอัตราประสิทธิภาพมากยิ่งขึ้น [6] เสนอการหาค่าควอนไทล์พารามิเตอร์โดยใช้โมเดลกำลังสอง (Quadratic Model) สำหรับมาตรฐาน MPEG-2 [7] เสนอการควบคุมอัตราสำหรับการสื่อสารบนช่องสัญญาณไร้สายเมื่อเกิดความผิดพลาดในการสื่อสาร โดยมีแนวความคิดใช้ข้อมูลป้อนกลับถึงสถานะของช่องสัญญาณเพื่อหาจำนวนบิตที่เหมาะสม โดยการหาค่าควอนไทล์พารามิเตอร์ใช้โปรแกรมพลวัตและตัวคูณลากรองจ์ แต่วิธีนี้มีความซับซ้อนสูงไม่เหมาะสมสำหรับการสื่อสารสัญญาณวีดิทัศน์แบบเวลาจริง [8] เสนอการควบคุมอัตราสำหรับการสื่อสารผ่านช่องสัญญาณที่ไม่คงที่ เช่น อินเทอร์เน็ต [9] เสนอการควบคุมอัตราสำหรับการสื่อสารบนช่องสัญญาณไร้สายเมื่อเกิดความผิดพลาดในการสื่อสาร โดยมีแนวความคิดการปรับตัวได้อัตโนมัติของตัวเข้ารหัส การกำหนดจำนวนของมาโครบล็อกสำหรับการเข้ารหัสภายในเฟรม และการปรับระดับค่าควอนไทล์พารามิเตอร์ [10][11] เสนอการควบคุมอัตราสำหรับการสื่อสารบนช่องสัญญาณไร้สายเมื่อเกิดความผิดพลาดในการสื่อสาร คือ การส่งซ้ำ (Retransmission) มีแนวความคิดการจำลองสถานะของช่องสัญญาณ และการป้อนกลับสถานะของช่องสัญญาณ เพื่อทำการหาจำนวนบิตที่เหมาะสม โดยใช้พื้นฐานของการควบคุมอัตราสำหรับมาตรฐาน H.263 [12] เสนอการควบคุมอัตราสำหรับการสื่อสารสัญญาณวีดิทัศน์แบบเวลาจริงสำหรับมาตรฐาน MPEG-4 เพื่อแก้ไขปัญหา การติดขัดของช่องสัญญาณ การสูญหายของข้อมูล โดยการส่งข้อมูลโดยแบ่งเป็นส่วน ๆ การทำสัญลักษณ์สำหรับการชิงโครโนซ์ และคำนวณค่าควอนไทล์พารามิเตอร์ที่เหมาะสมโดยสอดคล้องกับช่อง สัญญาณ

และจำนวนบิตสะสมในบัฟเฟอร์ [13] นำเสนอการควบคุมอัตราสำหรับมาตรฐาน MPEG-4 ปรับปรุงการหาบิตเป้าหมายให้มีประสิทธิภาพมากยิ่งขึ้นโดยการคำนึงถึงจำนวนบิตสะสมในบัฟเฟอร์

สำหรับการศึกษารวบรวมอัตราสำหรับมาตรฐานการลงรหัสวีดิทัศน์วีดิทัศน์ชั้นสูงได้มีการศึกษา แบ่งได้ 2 วิธี คือ 1.ใช้เทคนิคการควบคุมอัตราสำหรับ MPEG-2 TM5 มาใช้ในการควบคุม แต่เนื่องจากการประยุกต์ใช้งานสำหรับ MPEG-2 [15] เป็นการเก็บข้อมูลลงในสื่อเก็บข้อมูล ดังนั้นวิธีนี้จึงไม่เหมาะสำหรับการสื่อสารแบบเวลาจริง 2.การควบคุมอัตรา โดยใช้โมเดลกำลังสอง [12] [16] มาใช้ในการหาค่าควอนไทซ์พารามิเตอร์ และเสนอแนวคิดของหน่วยย่อยพื้นฐาน (basic unit) [17] เพื่อประยุกต์ใช้สำหรับการสื่อสารแบบเวลาจริง

ในการสื่อสารวีดิทัศน์เวลาจริง โดยเฉพาะอย่างยิ่งสำหรับการประยุกต์ ใช้งานแบบโต้ตอบสองทาง (two-way interactive) เช่น โทรศัพท์ภาพ (videophone) หรือ การประชุมสัมมนาทางวีดิทัศน์ การประวิงเวลาจากต้นทางถึงปลายทาง (end-to-end delay) ควรเป็นช่วงสั้น ๆ ถ้าจำนวนบิตที่เข้ารหัสแต่ละเฟรมมีขนาดมากกว่าจำนวนบิตที่คาดหมายไว้จำนวนบิตที่เกินจะถูกเก็บสะสมไว้ในบัฟเฟอร์ของตัวเข้ารหัส และเพิ่มการประวิงเวลาของบัฟเฟอร์ (buffer delay) ถ้าข้อมูลสะสมในบัฟเฟอร์มาก จะทำให้บัฟเฟอร์มากเกินเก็บ (buffer overflow) โดยปกติเพื่อลดการประวิงเวลาและจำนวนบิตสะสมในบัฟเฟอร์ ตัวเข้ารหัสจะทำการกระโดดข้ามเฟรม (frame skipping) ซึ่งเป็นสาเหตุของความไม่ต่อเนื่องด้านการเคลื่อนไหวของสัญญาณวีดิทัศน์

ในวิทยานิพนธ์ฉบับนี้จึงได้มีการศึกษาผลกระทบจากการสะสมของบัฟเฟอร์ของการควบคุมอัตราสำหรับมาตรฐานการลงรหัสวีดิทัศน์ชั้นสูงเมื่อจำกัดการประวิงเวลาพบว่าจำนวนเฟรมกระโดดมีค่ามากและคุณภาพสัญญาณวีดิทัศน์มีค่าลดลง จึงได้ศึกษาถึงปัจจัยที่มีผลต่อจำนวนบิตสะสมในบัฟเฟอร์ โดยการควบคุมอัตราที่นำเสนอแบ่งได้ 3 ระดับ คือ การควบคุมอัตราในระดับกลุ่มภาพ การควบคุมอัตราในระดับเฟรม และ การควบคุมอัตราในระดับหน่วยย่อยพื้นฐาน

1.2 มาตรฐานการลงรหัสวีดิทัศน์

มาตรฐานการลงรหัสวีดิทัศน์ได้รับการพัฒนาอย่างต่อเนื่องในช่วง 20 ปีที่ผ่านมาโดยมาตรฐานที่รู้จักกันอย่างแพร่หลายได้รับการจัดทำโดย 2 องค์กรหลัก ได้แก่ ITU-T และ ISO/IEC โดยได้กำหนดมาตรฐานการลงรหัสวีดิทัศน์ดังนี้

1.2.1 มาตรฐานของ ITU-T

องค์กร ITU-T เริ่มต้นตั้งแต่ปลายทศวรรษที่ 1980 มุ่งเน้นในการส่งข้อมูลวีดิทัศน์ที่ถูกบีบอัดผ่านโครงข่าย ใช้งานกับเครื่องโทรศัพท์ภาพ หรือ การประชุมสัมมนาทางวีดิทัศน์ โดยกำหนดมาตรฐานตระกูล H (H-series) มาตรฐานต่างๆ ในตระกูล H แสดงไว้ดังตารางที่ 1.1

ตารางที่ 1.1 มาตรฐานการลงรหัสสัญญาณวีดิทัศน์ของ ITU-T

หมายเลขมาตรฐาน	ปีที่กำหนดใช้	วัตถุประสงค์
H.261	ค.ศ.1990	เพื่อจัดการระบบการประชุมสัมมนาทางวีดิทัศน์ผ่านโครงข่าย ISDN (ISDN Videoconference network)
H.262 (ร่วมกับ ISO/IEC)	ค.ศ.1995	เพื่อการส่งกระจายสัญญาณ broadcast และพัฒนาการเก็บข้อมูลเหมือน MPEG-2
H.263	ค.ศ.1996	เพื่อจัดการและเข้ารหัสสัญญาณวีดิทัศน์เพื่อส่งผ่านโครงข่ายอัตราบิตต่ำและโครงข่ายไร้สาย
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	เพื่อจัดการและเข้ารหัสสัญญาณวีดิทัศน์เพื่อส่งผ่านโครงข่ายอินเทอร์เน็ตและการให้บริการสัญญาณวีดิทัศน์

1.2.2 มาตรฐานของ ISO/IEC

มาตรฐาน MPEG เป็นมาตรฐานซึ่งตั้งขึ้นโดยคณะกรรมการ MPEG (Moving Picture Experts Group) โดยเริ่มขึ้นในปี ค.ศ. 1988 โดยมีวัตถุประสงค์บีบอัดสัญญาณวีดิทัศน์

และออกดีโอ เพื่อเก็บลงสื่อดิจิทัล และส่งผ่านโครงข่ายอัตราบิตสูง โดยมาตรฐานต่างๆแสดงดังตารางที่ 1.2

ตารางที่ 1.2 มาตรฐานการลงรหัสสัญญาณวิดีโอของ ISO/IEC

หมายเลขมาตรฐาน	ปีที่กำหนดใช้	วัตถุประสงค์
MPEG-1	ค.ศ.1992	เพื่อลงรหัสข้อมูลวิดีโอสำหรับสื่อประสมและการเก็บข้อมูลในสื่อเก็บข้อมูล 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)	เพื่อการส่งสัญญาณวิดีโอแบบอัตราบิตต่ำมาก (Very low bit rate) ถึงสูงมาก และนำไปประยุกต์ใช้งานทางอินเทอร์เน็ตแบบโต้ตอบกันได้ (Interactive) มุ่งเน้นในการบีบอัดและใช้งานทางสื่อประสม
MPEG-7	ค.ศ. 2001	เพื่อพัฒนามาตรฐานการสืบค้นหาและการจัดเก็บข้อมูลในรูปแบบของสื่อประสม
MPEG-21	ค.ศ.2002	Multimedia Framework เพื่อวางระบบการสื่อสารสื่อประสมให้สามารถสื่อสารกันเป็นระบบเดียวกัน

สำหรับมาตรฐานการลงรหัสสัญญาณวิดีโอขั้นสูง เริ่มต้นเมื่อในปี ค.ศ. 1998 ได้ริเริ่มพัฒนามาตรฐานใหม่โดยเบื้องต้นใช้ชื่อว่า H.26L โดยต่อมาเปลี่ยนชื่อเป็น H.264 หรือ AVC โดยร่วมมือกันระหว่าง ITU-T และ ISO /IEC เป็น Joint video team (JVT) โดยคาดว่าจะออกเป็นมาตรฐานอย่างเป็นทางการก่อนสิ้นปี ค.ศ. 2003 วัตถุประสงค์เพื่อให้เหมาะสมสำหรับการประยุกต์ใช้กับงานทางวิดีโอแบบเต็มรูปแบบ[18] เช่น การสื่อสารสัญญาณวิดีโอที่

มีอัตราบิตต่ำสำหรับโครงข่ายไร้สาย การแพร่ภาพบนโทรทัศน์มาตรฐาน (standard-definition television) การแพร่ภาพบนโทรทัศน์ขั้นสูง (high definition broadcast television) การสตรีม (stream) สัญญาณวิดีโอผ่านอินเทอร์เน็ต (internet) และ การแสดงภาพคุณภาพสูงสุดสำหรับการประยุกต์ใช้ในโรงภาพยนตร์ เป็นต้น

1.3 แนวทางที่นำเสนอ

เนื่องจากในปัจจุบันงานวิจัยทางการลงรหัสสัญญาณวิดีโอ มีแนวโน้มของการพัฒนาการวิจัยทางการลงรหัสสัญญาณวิดีโอเพื่อใช้งานในด้านการสื่อสารที่อัตราบิตต่ำ ซึ่งการประยุกต์ใช้งานส่วนใหญ่ใช้สำหรับการสื่อสารเวลาจริงซึ่งการประวิงเวลาเป็นปัจจัยสำคัญ งานวิจัยที่ผ่านมาเป็นมาตรฐานการลงรหัสสัญญาณวิดีโอที่อัตราบิตต่ำ H.263 และ MPEG-4 [5] แต่ยังไม่มีการศึกษาในมาตรฐานการลงรหัสสัญญาณวิดีโอขั้นสูง

ดังนั้นในวิทยานิพนธ์นี้ จึงได้มีการศึกษาถึงผลกระทบของการควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวิดีโอขั้นสูงเมื่อจำกัดการประวิงเวลา แล้วพบว่าจำนวนเฟรมกระโดดส่งผลให้ลำดับภาพมีความไม่ต่อเนื่องและคุณภาพสัญญาณวิดีโอมีค่าลดลง ซึ่งไม่เหมาะสมสำหรับประยุกต์ใช้งานจริง จึงได้มีการปรับปรุงการเข้ารหัสโดยได้ศึกษาพารามิเตอร์ต่างๆ ที่มีผลต่อการเข้ารหัสสัญญาณวิดีโอโดยมุ่งเน้นที่การจัดการกับจำนวนบิตสะสมในบัพเฟอร์เพื่อป้องกันการกระโดดข้ามเฟรมในกรณีที่มีจำนวนบิตสะสมเกินจากค่าสูงสุดบัพเฟอร์

เพื่อป้องกันมิให้จำนวนบิตสะสมในบัพเฟอร์มากเกินไปหรือน้อยเกินไปจึงเสนอการปรับค่าถ่วงน้ำหนักบิตเป้าหมาย (Weighted Target bits: β) โดยปรับค่าอยู่ระหว่าง 0 ถึง 1 โดยการพัฒนาเทคนิคการปรับตัวค่าถ่วงน้ำหนักบิตเป้าหมายและเงื่อนไขสำหรับการเลือกค่าควอนไทซ์พารามิเตอร์จะคำนึงถึงจำนวนบิตสะสมในบัพเฟอร์ในขณะนั้น

1.4 วัตถุประสงค์

เพื่อปรับปรุงวิธีการเข้ารหัสสัญญาณวิดีโอตามมาตรฐานการลงรหัสสัญญาณวิดีโอขั้นสูงให้คุณภาพของสัญญาณวิดีโอดีขึ้นในการส่งแบบเวลาจริงภายใต้การประวิงเวลาต่ำ

1.5 ขอบเขตของวิทยานิพนธ์

ปรับปรุงขั้นตอนวิธีการควบคุมอัตราในระดับเฟรมและระดับหน่วยพื้นฐานย่อยของกระบวนการเข้ารหัสที่อัตราบิดต่ำตามมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูงในการส่งผ่านช่องสัญญาณที่มีการกำหนดการประวิงเวลาของตัวเข้ารหัสเท่ากับ 100 มิลลิวินาที และทำการวัดประสิทธิภาพของขั้นตอนวิธีที่นำเสนอในเชิงของจำนวนเฟรมกระโดดที่ลดลงและค่าคุณภาพสัญญาณวีดิทัศน์เพิ่มขึ้น

1.6 ขั้นตอนและวิธีการดำเนินงาน

1. ศึกษาความรู้พื้นฐานในการลงรหัสสัญญาณวีดิทัศน์และการส่งสัญญาณวีดิทัศน์
2. ศึกษาขั้นตอนวิธีการควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูง
3. ศึกษาปัจจัยที่มีผลต่อการประวิงเวลา
4. พัฒนาขั้นตอนวิธีการควบคุมอัตราในระดับเฟรมที่นำเสนอ
5. พัฒนาขั้นตอนวิธีการควบคุมอัตราในระดับหน่วยพื้นฐานย่อย
6. ทดสอบวิธีที่นำเสนอกับขั้นตอนวิธีการควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูง พร้อมทั้งวิเคราะห์ผล
7. สรุปผลการทดลองและเขียนวิทยานิพนธ์

1.7 ประโยชน์ที่คาดว่าจะได้รับ

1. ได้ศึกษาความรู้พื้นฐานในการลงรหัสสัญญาณวีดิทัศน์และการส่งสัญญาณวีดิทัศน์
2. พัฒนาซอฟต์แวร์อัลกอริทึมในส่วนของกระบวนการควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูงเพื่อใช้เป็นข้อมูลในการลงรหัสเพื่อให้ได้ประสิทธิภาพของสัญญาณวีดิทัศน์ที่ดีขึ้นในเชิงจำนวนเฟรมกระโดด และค่าคุณภาพสัญญาณวีดิทัศน์เมื่อเปรียบเทียบกับขั้นตอนวิธีการควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูงแบบไม่จำกัดการประวิงเวลา
3. สามารถนำไปประยุกต์ใช้งานกับการส่งสัญญาณวีดิทัศน์เวลาจริงหรือการประชุมสัมมนาทางวีดิทัศน์

บทที่ 2

ความรู้พื้นฐาน

เนื้อหาในบทนี้กล่าวถึงทฤษฎีต่างๆ ที่เกี่ยวข้องกับวิทยานิพนธ์ คือ พื้นฐานของวิดีโอทัศน์ดิจิทัล หลักการทำงานพื้นฐานของการลงทะเบียนสัญญาณวิดีโอทัศน์ การควบคุมอัตราสำหรับการลงทะเบียนสัญญาณวิดีโอทัศน์ขั้นสูงและการวัดคุณภาพของภาพ

2.1 วิดิทัศน์ดิจิทัล (Digital Video)

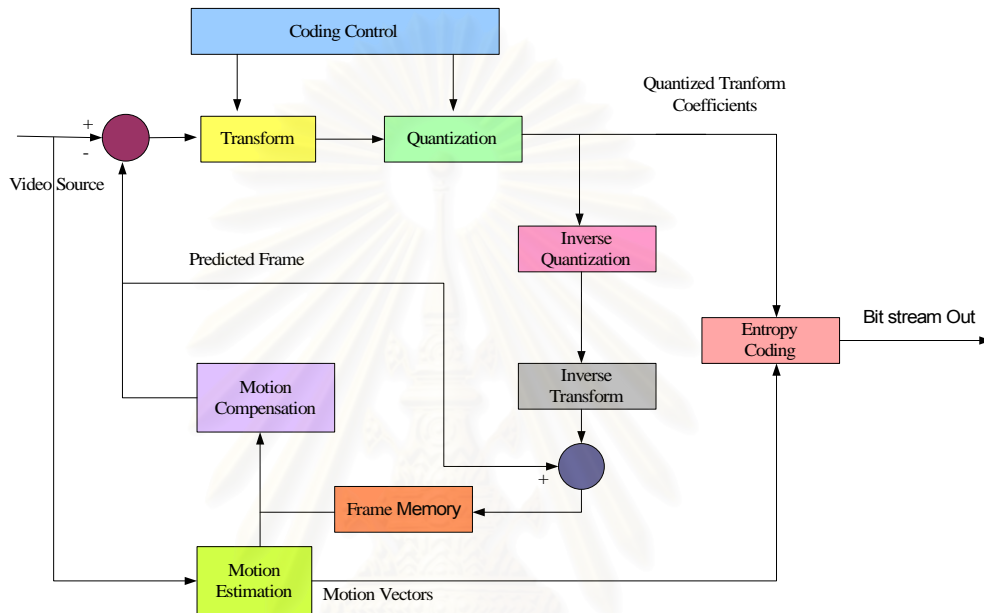
วิดีโอทัศน์ดิจิทัลเป็นกลุ่มภาพที่มากจากการบันทึกจากกล้องดิจิทัลหรือสแกนรูปจากภาพแอนาล็อก(analog) โดยมีหน่วยเป็นเฟรมต่อหนึ่งหน่วยเวลา เช่น เฟรมต่อวินาที แต่ละภาพจะประกอบไปด้วยจุดภาพมาจากการสุ่มตัวอย่างและผ่านการควอนไทซ์โดยแต่ละจุดภาพประกอบด้วย ส่วนประกอบความเข้มแสง (luminance component) และส่วนประกอบสัญญาณสี (chrominance component) โดยทั่วไป ความเข้มแสงของแต่ละภาพเมื่อผ่านการควอนไทซ์แล้วจะถูกแบ่งออกเป็น 256 ระดับ ดังนั้นจุดภาพแต่ละจุดสามารถแสดงด้วย ความเข้มแสงตั้งแต่ 0 ถึง 255 สำหรับความเข้มสีไม่มีความจำเป็นต้องแสดงทุกจุดภาพเนื่องจากความสามารถในการรับสีของระบบประสาทในการมองเห็นของมนุษย์น้อยกว่าความเข้มแสงจึงมีจำนวนจุดความเข้มสีแล้วแต่กำหนด [19] รูปแบบของสัญญาณวิดีโอทัศน์บางชนิดกำหนดไว้ดังตารางที่ 2.1

ตารางที่ 2.1 รูปแบบของสัญญาณวิดีโอทัศน์ที่ใช้งานในปัจจุบัน

รูปแบบ	ความละเอียด	อัตราเฟรม	Y:Cb:Cr	การกราด
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 การลงรหัสสัญญาณวิดีโอ (Video Coding)

ในปัจจุบันมีการใช้มาตรฐานการลงรหัสสัญญาณวิดีโอที่ต่างกันอย่างแพร่หลาย ซึ่งแต่ละมาตรฐานมีวัตถุประสงค์และการประยุกต์ใช้งานแตกต่างกันไป สำหรับการเข้ารหัสสัญญาณวิดีโอมีหลักการทำงานหลักๆแสดงดังรูปที่ 2.1



รูปที่ 2.1 บล็อกไดอะแกรมของการเข้ารหัสสัญญาณวิดีโอที่มีการบีบอัด

สัญญาณวิดีโอที่เข้าจะถูกแบ่งให้อยู่ในระดับบล็อกที่มีขนาดเท่ากับ 16x16 จุดภาพ หรือเรียกว่ามาโครบล็อก แต่ละมาโครบล็อกสามารถเข้ารหัสภายในเฟรม (Intra coding) หรือ การเข้ารหัสระหว่างเฟรม (Inter 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 จากนั้นสัมประสิทธิ์ที่ได้มาจะถูกนำจัดเรียงใหม่ให้อยู่ในรูปแบบเดียว โดยการกวาดแบบซิกแซก (zigzag scan) และทำการเข้ารหัสเอนโทรปี ในการเข้ารหัสสัญญาณวิดีโอ เมื่อผ่านการแปลง และการควอนไทซ์แล้วจะทำการแปลงย้อนกลับ เพื่อให้อยู่ในรูปแบบจุดภาพเช่นเดิมเพื่อนำไปประมาณการเคลื่อนที่ โดยตัวเข้ารหัสจะค้นหาบล็อกหน้าที่เก็บไว้ในหน่วยความจำว่าบล็อกใดที่มีค่าความแตกต่างระหว่างมาโครบล็อกของเฟรมปัจจุบันและเฟรมก่อนหน้า

ที่มีค่าน้อยที่สุด ทำให้ได้เวกเตอร์การเคลื่อนที่ (motion vector) เมื่อชดเชยการเคลื่อนที่ตัวเข้ารหัส จะเก็บสัญญาณวิดีโอที่เฟรมใหม่เพื่อสามารถใช้สำหรับเป็นเฟรมอ้างอิงต่อไป

ในการเข้ารหัสสัญญาณวิดีโอที่เฟรมสามารถแบ่งได้ 2 ประเภท คือ การเข้ารหัสภายในเฟรม เฟรมที่ถูกเข้ารหัสเรียกว่า 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

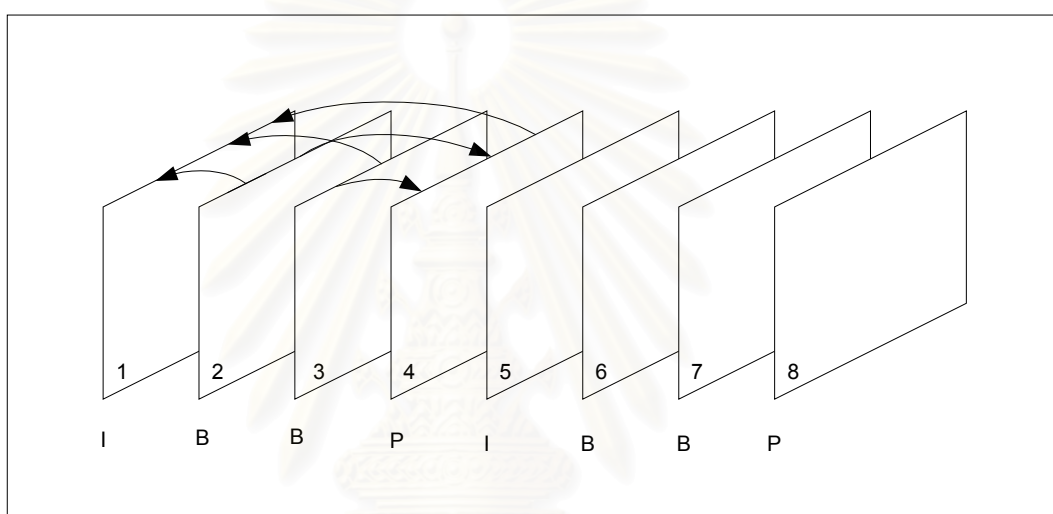
ชนิดของรูปภาพแบ่งตามลักษณะของการเข้ารหัสสามารถแบ่งได้ 3 ชนิด ได้แก่ I P และ B โดยมีรายละเอียด ดังนี้

1. I (Intra) เป็นภาพที่เข้ารหัสโดยไม่ต้องอาศัยข้อมูลจากเฟรมอื่น ดังนั้นลำดับภาพของวิดีโอที่เฟรมแรกจะเป็นภาพแบบ I เสมอ โดยเป็นเฟรมที่ถูกบีบอัดภายในเฟรม ใช้เป็นจุดเข้าถึง (access point) ของการเข้าถึงแบบสุ่ม ภาพแบบ I
2. P (Predicted) เป็นภาพที่เข้ารหัสโดยอาศัยข้อมูลของเฟรมก่อนหน้า และสามารถถอดรหัสโดยอาศัยข้อมูลของเฟรมก่อนหน้า ซึ่งเฟรมก่อนหน้าเรียกว่า เฟรมอ้างอิงซึ่งอาจจะเป็น I หรือ P เฟรมก็ได้ สำหรับการอ้างอิงถึงข้อมูลของเฟรมอ้างอิงต้องระบุเวกเตอร์การเคลื่อนที่มาด้วย ภาพแบบ
3. B (Bidirectionally interpolated) เป็นภาพที่เข้ารหัสและถอดรหัสโดยอาศัยข้อมูลทั้งจากอดีตและอนาคต รวมกับค่าการเคลื่อนที่ของเฟรมดังกล่าวทั้งสองเฟรม โดยละเลยค่าความเพี้ยนของภาพที่ได้จากการประมาณ ดังนั้นเฟรมชนิดนี้จะไม่ใช้เป็นเฟรมอ้างอิง

โดยถ้าเรียงจำนวนบิตที่ต้องใช้สำหรับเข้ารหัสชนิดต่างจากมากไปหาน้อย พบว่ามากที่สุดคือ ภาพแบบ I รองลงมาคือ ภาพแบบ P น้อยที่สุด คือ ภาพแบบ 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

2.2.2 การเข้ารหัสภายในเฟรม

การเข้ารหัสชนิดนี้เป็นการลดความซ้ำซ้อนในโดเมนปริภูมิ (Spatial Redundancy) ของเฟรมภาพด้วยการบีบอัดข้อมูลพื้นฐานการพิจารณาความซ้ำซ้อนของข้อมูลภายในหนึ่งเฟรม โดยเทคนิคต่างๆดังนี้



รูปที่ 2.2 ความสัมพันธ์ระหว่างเฟรมแบบ I P และ B ในการทำนายการเคลื่อนที่

1. การควอนไทซ์แบบหยาบหรือการลดความลึกของจุดภาพ (Coarse Quantization or Depth Reduction) เป็นการลดจำนวนข้อมูลภายในเฟรมโดยข้อมูลบางส่วนของภาพถูกทิ้งไป
2. การเข้ารหัสแบบการแปลง (Transform coding) ซึ่งเป็นการแปลงข้อมูลภาพจากโดเมนปริภูมิ (Spatial domain) เป็นโดเมนความถี่ (Frequency Domain) สัมประสิทธิ์จากการแปลงจะอยู่ในรูปสองมิติและถูกจัดให้อยู่มิติเดียวเพื่อทำการเข้ารหัสต่อไป
3. การทำควอนไทซ์แบบเวกเตอร์ (Vector Quantization) ข้อมูลภาพจะถูกแบ่งกลุ่มและแสดงแทนด้วยสัญลักษณ์จากชุดรหัส (code book) เพื่อทำการส่งสัญลักษณ์แทนเท่านั้น ซึ่งเป็นการประหยัดจำนวนบิตในการส่ง

สำหรับมาตรฐานการลงรหัสสัญญาณวิดีโอที่ศึ้นสูง [20] การเข้ารหัสภายในเฟรมแบ่งออกเป็นบล็อก ส่วนประกอบความเข้มแสงขนาด 4×4 จุดภาพทั้งหมด 9 โหมดและ บล็อกส่วนประกอบความเข้มแสงขนาด 16×16 จุดภาพ ทั้งหมด 4 โหมด และสำหรับบล็อกส่วนประกอบ

สี่ขนาด 4x4 จุดภาพ 4 โหมด โดยการเลือกโหมดของการเข้ารหัสเลือกจากการหาค่าความแตกต่างสัมบูรณ์รวม (Sum Absolute Difference :SAD) ที่น้อยที่สุด ดังสมการที่ (2.1)

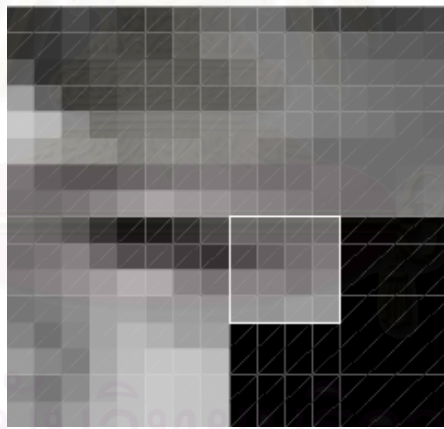
$$SAD = \sum_{i=1}^N \sum_{j=1}^N |I(i, j) - O(i, j)| \quad (2.1)$$

โดยที่ $N \times N$ เป็นขนาดบล็อก

$I(i, j)$ คือ ค่าส่วนประกอบความเข้มแสงที่ได้จากการประมาณ

$O(i, j)$ คือ ค่าส่วนประกอบความเข้มแสงดั้งเดิม

สำหรับบล็อกส่วนประกอบความเข้มแสงขนาด 4x4 จุดภาพที่จะทำการเข้ารหัสแสดงดังรูปที่ 2.3 โดยแต่ละจุดภาพ (a ถึง p) สามารถแสดงดังรูปที่ 2.4 โดยแต่ละโหมดมีรายละเอียดดังนี้



รูปที่ 2.3 ส่วนประกอบความเข้มแสงขนาด 4 x 4 จุดภาพ [20]

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 ตำแหน่งส่วนประกอบความเข้มแสงขนาด 4 x 4 จุดภาพ

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

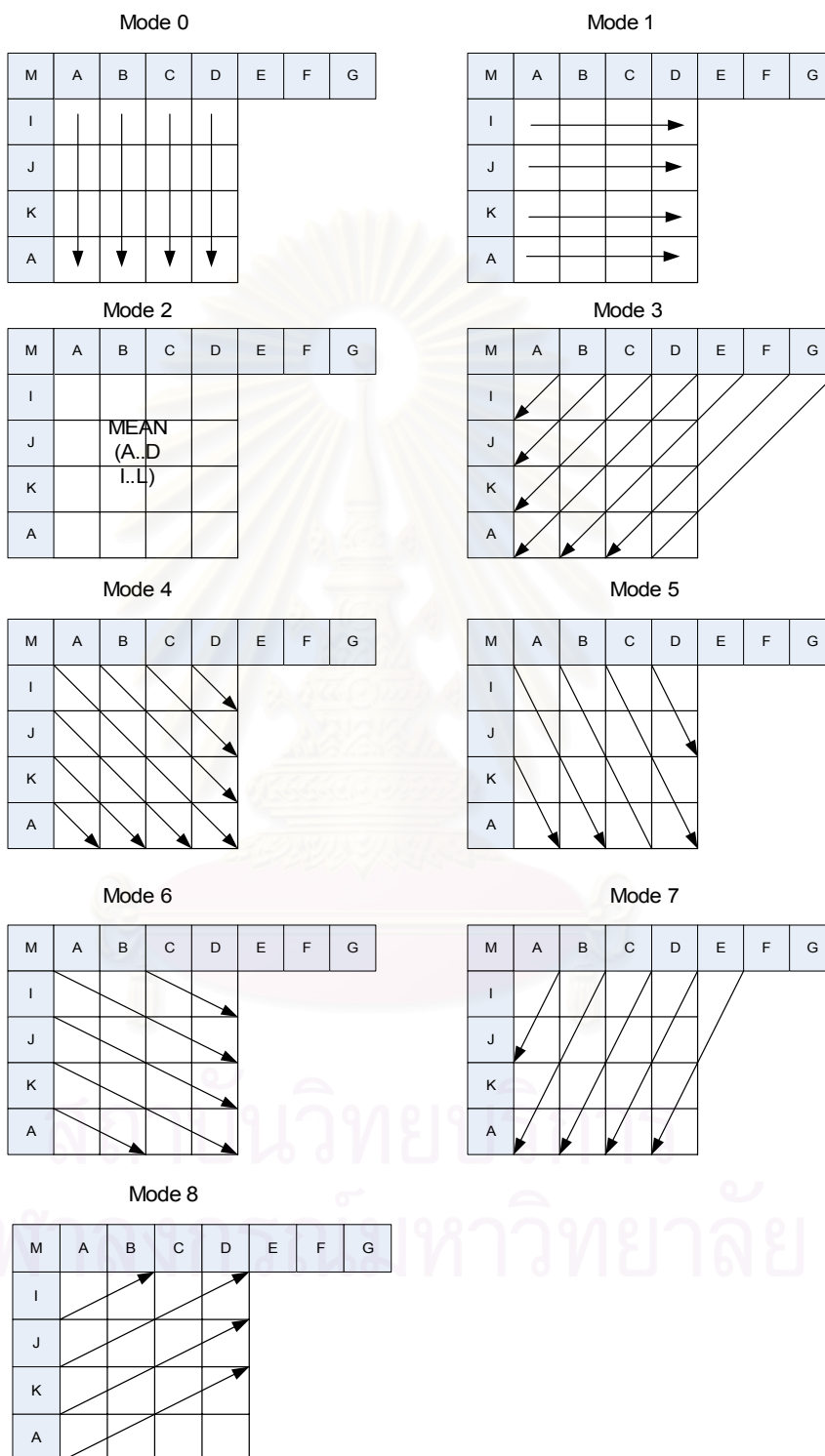
โดยสามารถแสดงดังรูปที่ 2.5

สำหรับบล็อกส่วนประกอบความเข้มแสงขนาด 16x16 จุดภาพ สามารถหาได้

จาก 4 โหมต รายละเอียดดังนี้

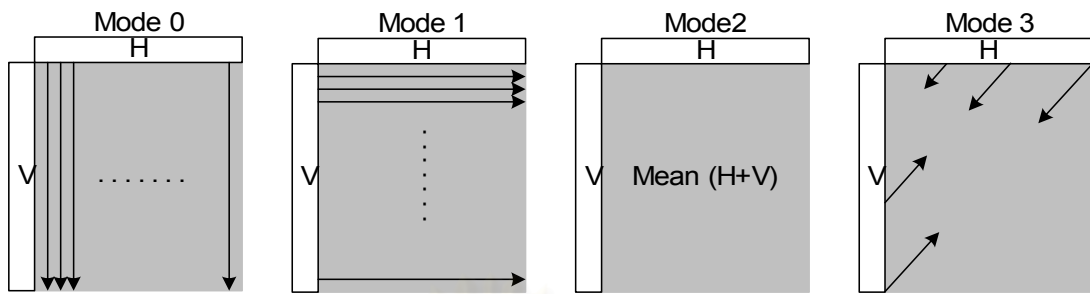
- โหมต 0 สามารถคำนวณจากการประมาณในแนวตั้งจากกลุ่มจุดภาพ H
- โหมต 1 สามารถคำนวณจากการประมาณในแนวตั้งจากกลุ่มจุดภาพ V
- โหมต 2 สามารถคำนวณจากค่าเฉลี่ยจากกลุ่มจุดภาพ H และ V
- โหมต 3 สามารถคำนวณจากกลุ่มจุดภาพ H ในแนวทแยงจากขวามาซ้าย และกลุ่มจากจุดภาพ V ในแนวทแยงจากซ้ายไปขวา

โดยสามารถแสดงดังรูปที่ 2.6



รูปที่ 2.5 การเข้ารหัสภายในเฟรมของส่วนประกอบความเข้มแสงขนาด 4x4 จุดภาพ

โหมด 0 ถึง 8



รูปที่ 2.6 การเข้ารหัสภายในเฟรมของส่วนประกอบความเข้มแสงขนาด 16x16 จุดภาพ โหมด 0 ถึง 3

สำหรับการเข้ารหัสภายในเฟรมของส่วนประกอบความเข้มสีขนาด 8x8 จุดภาพ มีทั้งหมด 4 โหมดลักษณะคล้ายกับการเข้ารหัสภายในเฟรมของส่วนประกอบความเข้มแสงขนาด 16x16 จุดภาพ

2.2.3 การเข้ารหัสระหว่างเฟรม

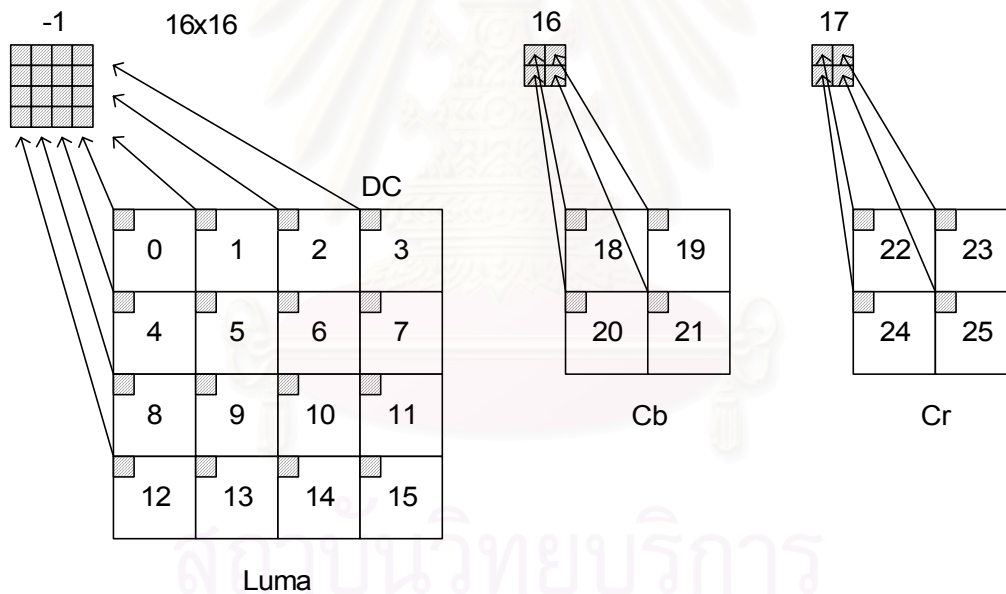
เนื่องจากสัญญาณวิดีโอที่ติดกันจะมีความสัมพันธ์ระหว่างข้อมูลทางปริภูมิ-เวลา (Spatial temporal correlation) ซึ่งทำให้เกิดความซ้ำซ้อนของข้อมูล เพื่อต้องการลดข้อมูลที่ซ้ำซ้อนนี้จึงใช้เทคนิคต่าง ๆ ดังนี้

1. การเข้ารหัสแบบซีกตัวอย่างย่อยทางเวลา (Time Sub-sampling Coding) การเข้ารหัสเฉพาะบางเฟรม วิธีนี้ทำให้ลดข้อมูลได้มากแต่อาจทำให้ลำดับภาพขาดความต่อเนื่อง
2. การเข้ารหัสค่าความต่าง (Different Coding) มี 2 แบบ คือ การเข้ารหัสเฉพาะความต่างระหว่างเฟรม และการเข้ารหัสค่าความต่างระหว่างบล็อก (Block-Based Different Coding) โดยการเข้ารหัสเฉพาะความต่างระหว่างเฟรมอาศัยความสัมพันธ์ของลำดับภาพที่อยู่ติดกันมีความคล้ายคลึงกัน แต่อาจต้องใช้ส่วนหัว (overhead) ในการบ่งบอกว่าจุดภาพใดมีการเปลี่ยนแปลง จุดภาพใดไม่มีการเปลี่ยนแปลง ส่วนการเข้ารหัสค่าความต่างระหว่างบล็อกต่างกันตรงที่ภาพจะถูกแบ่งเป็นบล็อกย่อย ๆ ก่อน หลังจากนั้นแต่ละบล็อกจะถูกนำมาเปรียบเทียบ
3. การประมาณการเคลื่อนที่ (Motion Estimation) และ การชดเชยการเคลื่อนที่ (Motion Compensation) เป็นวิธีการลดข้อมูลในการเข้ารหัส โดยเข้ารหัสเฉพาะความแตกต่างแทนที่จะเข้ารหัสทั้งเฟรม โดยทั่วไป เราสามารถสมมุติว่าค่าของจุดภาพในภาพที่มีการเคลื่อนที่ของวัตถุได้เลื่อนจากจุดหนึ่งในเฟรมหนึ่งไปยังจุดอื่นในเฟรมถัดไป ถ้าเราสามารถ

หาตำแหน่งที่เปลี่ยนแปลงหรือที่เรียกว่า เวกเตอร์การเคลื่อนที่ ทำให้สามารถลดความแตกต่างลงไปได้ วิธีนี้เรียกว่า การชดเชยการเคลื่อนที่ การประมาณการเคลื่อนที่ซึ่งมีบทบาทสำคัญในส่วนของ 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 ซึ่งกรณีที่ทำอินเวอร์สการแปลงจะมีประสิทธิภาพที่ดีกว่ารูปแบบการแปลงแบบ DCT โดยถ้าบล็อกนั้นเข้ารหัสแบบภายในเฟรม ค่าสัมประสิทธิ์กระแสตรง DC (Direct Coefficient) ของบล็อกส่วนประกอบความเข้มแสงจะถูกนำมาจัดเรียงและทำการส่งไปก่อน แล้วจึงทำการส่งข้อมูลของบล็อกที่เหลือแสดงดังรูปที่ 2.7



รูปที่ 2.7 การจัดเรียงค่าสัมประสิทธิ์

ภาพขาเข้า X ขนาด 4×4 การแปลงสามารถคำนวณได้จากสมการที่ (2.2)

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

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

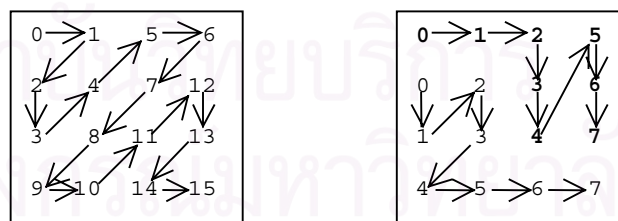
เมื่อ

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

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

2.2.5 การควอนไทซ์และการสแกนค่าสัมประสิทธิ์ที่ได้

การควอนไทซ์เป็นขั้นตอนของการสูญเสียข้อมูลบางส่วน การสูญเสียข้อมูลขึ้นอยู่กับค่าควอนไทซ์ กล่าวคือถ้าค่าควอนไทซ์มีค่ามากสูญเสียข้อมูลไปมาก ค่าควอนไทซ์มีค่าน้อยสูญเสียข้อมูลไปน้อย ทั้งนี้การเลือกค่าควอนไทซ์เป็นขั้นตอนที่มีความสำคัญต่อจำนวนบิตที่ทำการเข้ารหัสเฟรมนั้นๆ 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) และเมื่อผ่านการควอนไทซ์ เรียงค่าสัมประสิทธิ์ให้เป็นมิติเดียวโดยเริ่มจากค่าสัมประสิทธิ์ DC ซึ่งในมาตรฐาน H.264 นี้จะมีวิธีเรียงค่าสัมประสิทธิ์อยู่ 2 วิธีคือ การกราดแบบซิกแซก (Zigzag scan) เป็นวิธีที่ใช้โดยทั่วไปในมาตรฐาน H.263 วิธีที่ 2 คือ การกราดแบบคู่ (Double scan) ลักษณะการเรียงค่าสัมประสิทธิ์ทั้ง 2 วิธี แสดงดังรูปที่ 2.8



รูปที่ 2.8 วิธีกราดค่าสัมประสิทธิ์แบบซิกแซกและแบบคู่

2.2.6 การเข้ารหัสเอ็นโทรปี

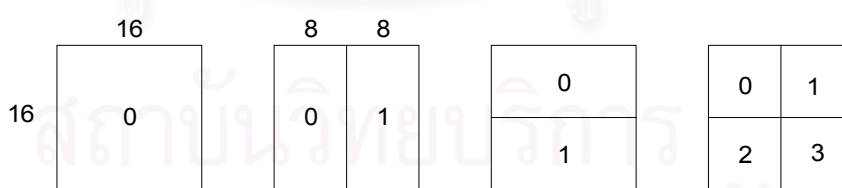
เป็นขั้นตอนการแปลงค่าจากสัมประสิทธิ์ทางความถี่ที่ได้จากการควอนไทซ์ให้เป็นเลขไบนารีเพื่อส่งเป็นบิตสตรีมออกไป โดยมาตรฐาน H.264 มีวิธีการเข้ารหัสเอ็นโทรปี 2 วิธี

คือ Context-Adaptive Variable Length Coding (CAVLC) [21] และ Context-Based Adaptive Arithmetic Coding (CABAC) [21] โดยวิธี CABAC สามารถลดอัตราบิตได้มากกว่า CAVLC ประมาณ 5%-15% [18]

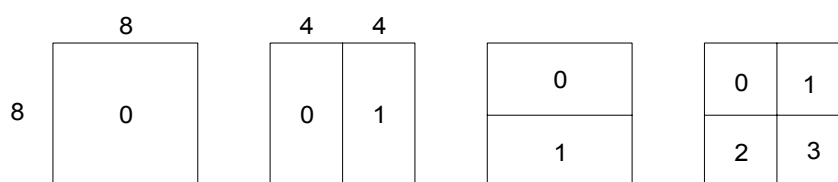
2.2.7 การประมาณการเคลื่อนที่ (Motion Estimation) และ การชดเชยการเคลื่อนที่ (Motion compensation)

การประมาณการเคลื่อนที่เป็นการหาความแตกต่างระหว่างจุดภาพที่สอดคล้องกันระหว่างเฟรมสองเฟรมที่ติดกันโดยปกติเฟรมที่อยู่ติดกันจะมีค่าสหสัมพันธ์สูงหรือที่เรียกว่า ความซ้ำซ้อนทางเวลา ส่งผลให้สามารถเข้ารหัสเฉพาะความแตกต่างแทนที่จะเข้ารหัสทั้งภาพ โดยทั่วไป เราสามารถสมมุติว่าค่าของจุดภาพในภาพที่มีการเคลื่อนที่ของวัตถุได้เลื่อนจากจุดหนึ่งในเฟรมหนึ่งไปยังจุดอื่นในเฟรมถัดไป ถ้าเราสามารถหาตำแหน่งที่เปลี่ยนแปลงหรือที่เรียกว่า เวกเตอร์การเคลื่อนที่ ทำให้สามารถลดความแตกต่างลงไปได้ วิธีนี้เรียกว่า การชดเชยการเคลื่อนที่

โดยทั่วไปขนาดของบล็อกที่ใช้ในการชดเชยการเคลื่อนที่มีขนาด 16x16 หรือ 8x8 จุดภาพ ซึ่งในมาตรฐาน H.264 สำหรับบล็อกที่มีขนาด 16x16 จุดภาพ สามารถแบ่งขนาดบล็อกให้มีขนาดเล็กได้ทั้งหมด 4 แบบ คือ 16x16 16x8 8x16 8x8 จุดภาพ แสดงดังรูปที่ 2.9 และสำหรับบล็อกที่มีขนาด 8x8 จุดภาพ สามารถแบ่งให้มีขนาดเล็กได้อีก 4 แบบ คือ 8x8 8x4 4x8 และ 4x4 จุดภาพ แสดงดังรูป 2.10



รูปที่ 2.9 การแบ่งบล็อกขนาด 16x16 จุดภาพ



รูปที่ 2.10 การแบ่งบล็อกขนาด 8x8 จุดภาพ

สำหรับการเลือกขนาดของบล็อกโดยทั่วไป บล็อกขนาดใหญ่เหมาะสำหรับพื้นที่ที่มีลักษณะเอกพันธ์ (Homogeneous) และบล็อกขนาดเล็กเหมาะสำหรับพื้นที่ที่มีรายละเอียดมาก การแบ่งบล็อกแสดงดังรูปที่ 2.11[20]



รูปที่ 2.11 แสดงการแบ่งบล็อกภายในภาพ[20]

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

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

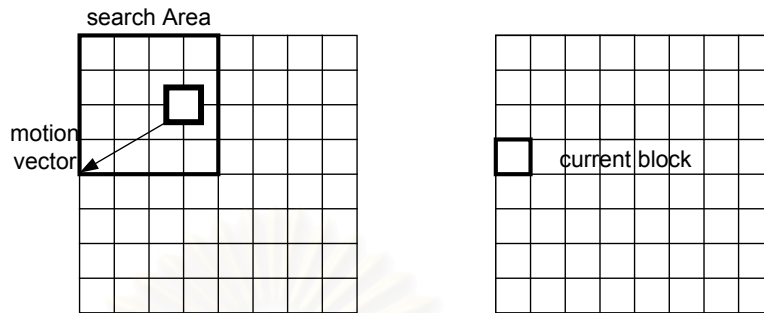
$$SSD(dx, dy) = \frac{1}{N \times N} \sum_{m=x}^{x+N-1} \sum_{n=y}^{y+N-1} (I_k(m, n) - I_k(m + dx, n + dy))^2 \quad (2.4)$$

โดยที่ $N \times N$ คือ ขนาดของบล็อก

$I_k(m, n)$ คือ ค่าความเข้มแสงของเฟรม k ที่ตำแหน่ง (m, n)

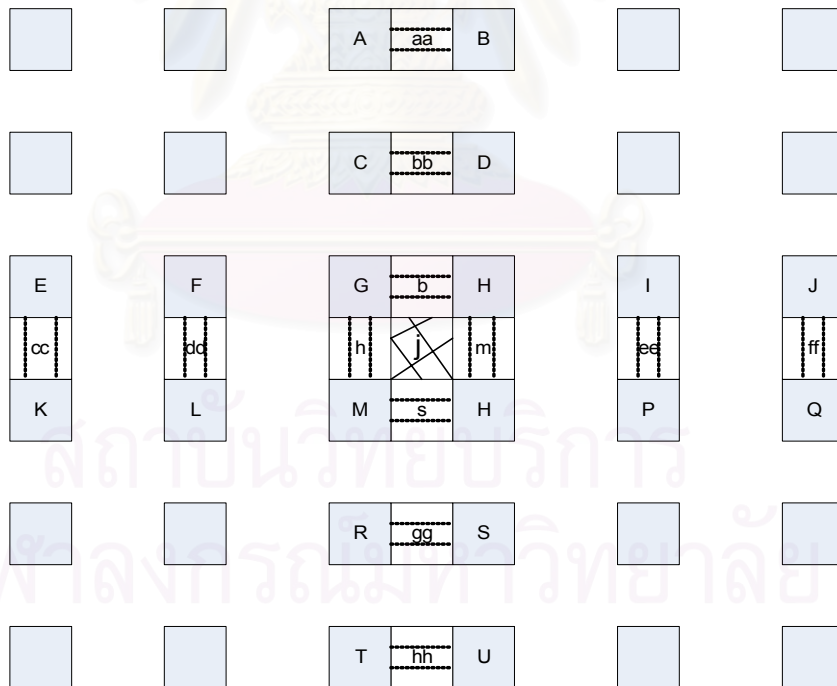
$MAD(dx, dy)$ คือ ค่าความแตกต่างสัมบูรณ์เฉลี่ย เมื่อเวกเตอร์การเคลื่อนที่เท่ากับ (m, n)

$SSD(dx, dy)$ คือ ค่าความแตกต่างกำลังสองรวม เมื่อเวกเตอร์การเคลื่อนที่เท่ากับ (m,n)



รูปที่ 2.12 การประมาณการเคลื่อนที่

สำหรับมาตรฐาน H.264 ได้เพิ่มระดับความแม่นยำของการประมาณการเคลื่อนที่ ไหวจากเดิมแบบครึ่งจุดภาพ (Half-pixel Motion Estimation) เป็น 1/4 และ 1/8 จุดภาพซึ่งสามารถทำให้ประมาณการเคลื่อนที่ได้แม่นยำยิ่งขึ้น สำหรับการประมาณการเคลื่อนที่ ไหวจากเดิมแบบครึ่งจุดภาพ สามารถแสดงได้ดังรูป 2.13



รูปที่ 2.13 การหาค่าจุดภาพระหว่าง 2 จุดภาพ

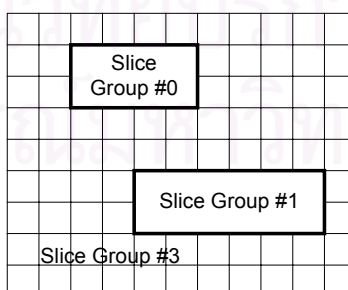
การหาค่าระหว่างจุดภาพสามารถหาได้จากการประมาณค่าในช่วง (Interpolated) โดยใช้วงจรกรองตอบสนองอิมพัลส์จำกัด (Finite Impulse Response (FIR) filter) 6 แถบ ซึ่งวง

จรรจงมีค่าถ่วงน้ำหนักคือ $(1/32 \ -5/32 \ 5/8 \ 5/8 \ -5/32 \ 1/32)$ [20] เช่น ต้องการหาจุดภาพ b สามารถคำนวณได้จากจุดภาพ 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 ได้มีการเพิ่มโหมดการทำงานเพื่อการบีบอัดสัญญาณวิดีโอที่ทันสมัยขึ้น [18] ดังนี้

1. การคำนวณเวกเตอร์การเคลื่อนที่แบบไม่จำกัด (Unrestricted Motion Vector Mode) นั้นคือ ถ้าเวกเตอร์การเคลื่อนที่ไหวซึ่งออกนอกพื้นที่ภาพ จะใช้วิธีการประมาณนอกช่วง (extrapolation) ซึ่งเทคนิคนี้อยู่ในโหมดพิเศษในมาตรฐาน H.263
2. การทำนายด้วยค่าถ่วงน้ำหนัก (Weighted prediction) การชดเชยการเคลื่อนที่ที่สามารถถ่วงน้ำหนักบริเวณที่สนใจ
3. การถอดคู่ควบลำดับเฟรมอ้างอิงจากลำดับการแสดงผลภาพ (Decoupling of reference)
4. สามารถอ้างอิงการประมาณการเคลื่อนที่และการชดเชยการเคลื่อนที่สูงสุด 5 เฟรมก่อนหน้าทำให้สามารถประมาณการเคลื่อนที่ได้แม่นยำมากขึ้น แต่ต้องคำนึงถึงหน่วยความจำที่ใช้เก็บรายละเอียดของเฟรมก่อนหน้า โดยเทคนิคนี้ใช้กับมาตรฐาน H.263++ เช่นกัน
5. สามารถจัดแบ่งกลุ่มของมาโครบล็อกโดยเรียกว่า การเรียงลำดับมาโครบล็อกแบบยืดหยุ่น (Flexible Macroblock Ordering) ประโยชน์ของการเรียงลำดับมาโครบล็อกแบบยืดหยุ่นนี้สำหรับการแบ่งบริเวณที่สนใจ (Region of Interest) แสดงดังรูปที่ 2.14



รูปที่ 2.14 การจัดกลุ่มมาโครบล็อกแบบยืดหยุ่น

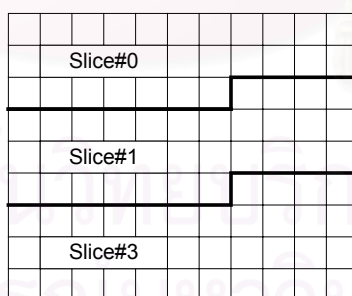
6. สามารถเข้ากับโครงข่ายต่าง ๆ ได้ง่าย(Network friendly)

7. สามารถแบ่งข้อมูลออกเป็นส่วน ๆ (Data Partitioning) โดยถ้าข้อมูลส่วนนั้นมีความสำคัญจะทำการส่งข้อมูลส่วนนั้นก่อน โดยเทคนิคใช้กับมาตรฐาน H.263++ เช่นกัน

2.2.8 รูปแบบการจัดเรียงสัญญาณ

ในมาตรฐาน H.264 ได้กำหนดรูปแบบการจัดเรียงข้อมูลในสัญญาณที่ถูกเข้ารหัสตามลำดับชั้นต่าง ๆ ทั้งนี้การจัดเรียงข้อมูลรวมทั้งข้อมูลในส่วนหัวของสัญญาณต้องเป็นไปตามมาตรฐาน เพื่อให้ไม่ให้เกิดข้อผิดพลาดจากความสับสน โดยจะแบ่งเป็นลำดับชั้นต่าง ๆ ได้ดังนี้

- 1) ชั้นกลุ่มภาพ (Group of Picture Layer) ประกอบด้วยจำนวนเฟรมหลายๆเฟรมอยู่ในกลุ่มเดียวกัน และแต่ละกลุ่มภาพจะเริ่มต้นด้วยภาพแบบ I มีเพียงเฟรมเดียวเท่านั้น
- 2) ชั้นรูปภาพ (Picture Layer) ประกอบด้วยจุดภาพขนาดเท่ากับสัญญาณวีดิทัศน์ โดยทั่วไปใน 1 เฟรมจะประกอบด้วย 2 ส่วนคือ บรรทัดคู่และบรรทัดคี่
- 3) ชั้นสไลซ์และกลุ่มสไลซ์ (Slices and Slice Groups) โดยแต่ละสไลซ์จะประกอบด้วยกลุ่มของมาโครบล็อกซึ่งได้มาจากการกวาดแบบราสเตอร์ (raster scan) การแบ่งกลุ่มสไลซ์แสดงดังรูปที่ 2.15



รูปที่ 2.15 การเรียงลำดับมาโครบล็อกแบบปกติ

- 4) ชั้นมาโครบล็อก (Macroblock layer) แต่ละมาโครบล็อกจะประกอบไปด้วยจุดภาพขนาด 16 x 16 จุดภาพ
- 5) ชั้นบล็อก (Block layer) ประกอบไปด้วยจุดภาพขนาด 8 x 8 จุดภาพ

2.2.9 รูปแบบสัญญาณวีดิทัศน์ที่ใช้

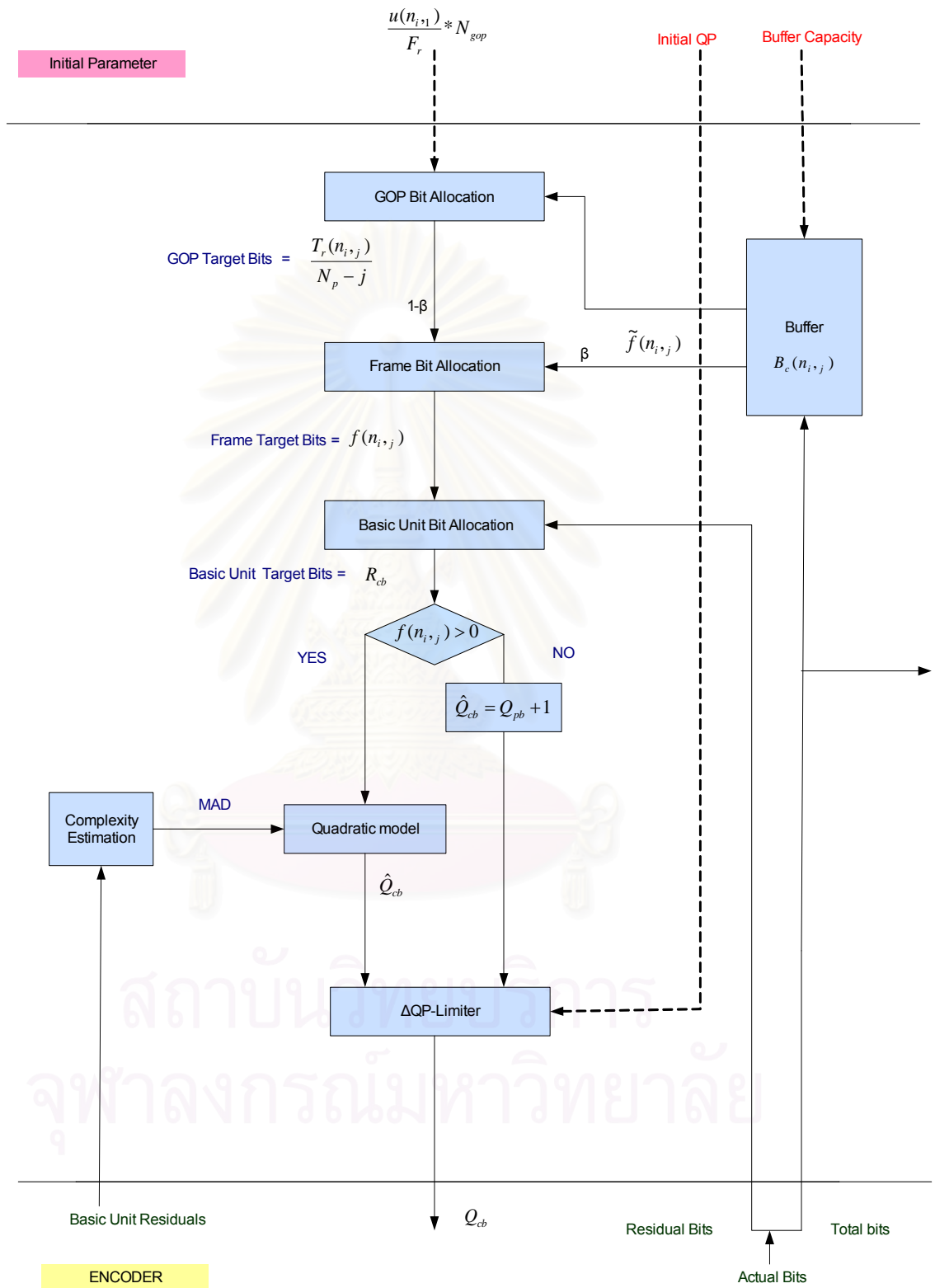
รูปแบบที่ใช้ในมาตรฐาน H.263 คือ CIF (Common Intermediate Format) และ QCIF (Quarter CIF) ซึ่งมีความละเอียดลดลงครึ่งหนึ่งจากรูปแบบ CIF แต่ละรูปแบบประกอบด้วยส่วนประกอบความเข้มสีและส่วนประกอบความเข้มแสง โดยรูปแบบ CIF จะมีส่วนประกอบความเข้มแสง Y มีจำนวนจุดภาพในแนวนอนเป็น 352 จุดภาพ และในแนวตั้ง 288 จุดภาพ ส่วนในรูปแบบ QCIF จะมีความละเอียดลดลงครึ่งหนึ่ง นั่นคือ จำนวนจุดภาพแนวนอนเป็น 176 จุดภาพ และในแนวตั้ง 144 จุดภาพ ในส่วนประกอบความเข้มสีของรูปแบบ CIF และ QCIF (Cb และ Cr) จะมีความละเอียดน้อยกว่าส่วนประกอบความเข้มแสง เนื่องจากตาคนไวต่อความเข้มแสงแต่ไม่ไวต่อความเข้มสี จึงไม่จำเป็นต้องใช้ข้อมูลของความเข้มสีละเอียดมากนัก

2.3 การควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูง [17]

การควบคุมอัตราสำหรับ AVC แบ่งออกได้ 3 ระดับ คือ ระดับกลุ่มภาพ (GOP level) ระดับเฟรม (Frame level) และระดับพื้นฐานย่อย (Basic unit level) สามารถแสดงดังแสดงในรูปที่ 2.16 โดยที่ในระดับกลุ่มภาพทำการคำนวณจำนวนบิตสำหรับเข้ารหัสกลุ่มภาพนั้น ๆ จากนั้นในระดับเฟรมภาพ (Frame level) ทำการคำนวณหาจำนวนบิตเป้าหมายที่ทำการเข้ารหัสในแต่ละเฟรม ส่วนในระดับหน่วยพื้นฐานย่อย (Basic unit level) จะทำการหาจำนวนบิตเป้าหมายในหน่วยพื้นฐานย่อยนั้น ๆ เมื่อได้จำนวนบิตเป้าหมายในระดับหน่วยพื้นฐานย่อยแล้วจะทำการหาค่าควอนไทซ์พารามิเตอร์โดยใช้โมเดลกำลังสอง ซึ่งการหาค่าควอนไทซ์พารามิเตอร์นี้ใช้ค่าความแตกต่างสัมบูรณ์ (Mean Absolute Difference) ด้วย จากนั้นทำการตรวจสอบเงื่อนไขเพื่อกำหนดค่าควอนไทซ์พารามิเตอร์ เมื่อเข้ารหัสครบทุกหน่วยพื้นฐานย่อย ทำการรวมจำนวนบิตทั้งหมดภายในเฟรม แล้วทำการคำนวณหาจำนวนบิตสะสมคงเหลือสะสมในบัฟเฟอร์ โดยบล็อกไดอะแกรมของการควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูง แสดงดังรูปที่ 2.16

2.3.1 การควบคุมอัตราระดับกลุ่มภาพ

ในระดับนี้คำนวณหาจำนวนบิตสำหรับเข้ารหัสภาพภายในกลุ่มภาพ ในกรณีที่เป็นเฟรมแรกของกลุ่มภาพ จำนวนบิตทั้งหมดคำนวณดังสมการที่ (2.6)



รูปที่ 2.16 บล็อกไดอะแกรมการควบคุมอัตราสำหรับมาตรฐานการลงรหัส
สัญญาณวิดีโอที่ซับซ้อนสูง

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

โดยที่

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

T_r คือ จำนวนบิตสำหรับกลุ่มภาพ

N_{gop} คือ จำนวนเฟรมในแต่ละกลุ่มภาพ (GOP)

F_r คือ อัตราเฟรม

B_s คือ ขนาดบัพเฟอร์เริ่มต้นโดยเท่ากับสองเท่าของอัตราบิตในช่องสัญญาณ

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

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

โดยจำนวนบิตสะสมในบัพเฟอร์สามารถคำนวณโดยใช้โมเดลทราฟฟิกการไหลของของไหล โดยจำนวนบิตสะสมในบัพเฟอร์หลังจากเข้ารหัสเฟรม j ($B_c(n_{i,j+1})$) หาได้จากการเลือกค่าน้อยที่สุดระหว่างค่าบัพเฟอร์เริ่มต้นกับค่ามากที่สุดระหว่างค่าศูนย์กับจำนวนบิตสะสมในบัพเฟอร์ของเฟรมก่อนหน้ากับผลต่างระหว่างจำนวนบิตที่ใช้ในการเข้ารหัสเฟรมก่อนหน้ากับจำนวนบิตที่สามารถออกสู่ช่องสัญญาณได้ แสดงในสมการที่ (2.7)

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

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

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

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

ดังนั้นถ้าทำการเข้ารหัสแต่ละเฟรมเสร็จต้องทำการคำนวณจำนวนบิตคงเหลือสามารถคำนวณดังแสดงในสมการที่ (2.8)

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

2.3.2 การควบคุมอัตราระดับเฟรม

ในระดับนี้แบ่งการทำงานได้ เป็น 2 ขั้นตอน

1. Pre-Encoding Stage ทำการคำนวณหาจำนวนบิตเป้าหมายเพื่อเข้ารหัสในแต่ละเฟรม แต่สำหรับกรณีที่เข้ารหัสเพียง 2 ระดับ คือ ระดับกลุ่มภาพ และ ระดับเฟรม ในขั้นตอนนี้จะคำนวณค่าควอนไทซ์พารามิเตอร์ด้วย
2. Post-Encoding Stage ทำการปรับค่าพารามิเตอร์ต่าง ๆ หลังการเข้ารหัสแต่ละเฟรม

2.3.2.1 ส่วน Pre-Encoding Stage

ในขั้นตอนนี้คำนวณหาจำนวนบิตเป้าหมายในแต่ละเฟรมและการคำนวณค่าควอนไทซ์พารามิเตอร์โดยมี 2 ขั้นตอนดังนี้

ขั้นตอนที่ 1 กำหนดบิตเป้าหมาย (Target bits) สำหรับเฟรมภาพ โดยบิตเป้าหมายสามารถคำนวณจาก 2 ส่วน คือ ส่วนที่มาจากระดับกลุ่มภาพและส่วนที่มาจากระดับบัฟเฟอร์ ขั้นตอนนี้แบ่งได้ 2 ขั้นตอนย่อยดังนี้

ขั้นตอนที่ 1.1 การควบคุมมหรรศน์ (Macroscopic Control)

การทำนายระดับบัฟเฟอร์เป้าหมาย (Target Buffer level) เพื่อนำไปใช้ในการคำนวณบิตเป้าหมายแสดงในสมการที่ (2.9)

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

เมื่อ $Tbl(n_{i,j})$ คือ ระดับบัฟเฟอร์เป้าหมายในเฟรมที่ j ในกลุ่มภาพที่ i
 N_p คือ จำนวนเฟรมภาพ P ในกลุ่มภาพ

ขั้นตอนที่ 1.2 การควบคุมจุลทรรศน์ (Microscopic control)

จากที่กล่าวข้างต้นการคำนวณบิตเป้าหมาย $f(n_{i,j})$ ประกอบด้วย 2 ส่วน ส่วนที่ใช้ระดับบัฟเฟอร์ ($\tilde{f}(n_{i,j})$) 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 คือ ค่าคงที่ ในที่นี้กำหนดให้มีค่าเท่ากับ 0.75

ส่วนที่มาจากระดับกลุ่มภาพหาได้จากอัตราส่วนระหว่างจำนวนบิตคงเหลือสำหรับเข้ารหัสเฟรมในกลุ่มภาพและจำนวน P เฟรมคงเหลือ แล้วทำการถ่วงน้ำหนักบิตเป้าหมายเพื่อให้ค่ามีความถูกต้องมากขึ้นดังแสดงในสมการที่ (2.11)

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

เมื่อ β คือ ค่าถ่วงน้ำหนักมีค่าตั้งแต่ 0 ถึง 1 ในที่นี้กำหนดให้มีค่าเท่ากับ 0.5
ขั้นตอนที่ 2 คำนวณค่าควอนไทซ์พารามิเตอร์โดยใช้โมเดลกำลังสอง [12] [16] [22] ดังแสดงในสมการที่ (2.12)

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

เมื่อ R_i คือ จำนวนบิตที่จะเข้ารหัสเฟรม i Q_i คือ ค่าควอนไทซ์สำหรับเฟรม i b_1 และ b_2 คือ ค่าสัมประสิทธิ์อันดับที่ 1 และอันดับที่ 2 ตามลำดับ

การหาค่าควอนไทซ์สำหรับโมเดลกำลังสอง ต้องพิจารณาอีก 2 ปัจจัยคือ

1. โมเดลกำลังสองพัฒนาจากสมมุติฐานที่ว่าเฟรมแต่ละเฟรมที่อยู่ติดกันสามารถเข้ารหัสด้วยความซับซ้อนที่คล้ายกัน โดยไม่คำนึงถึงความซับซ้อนของแต่ละเฟรม จึงต้องมีดัชนี (index) ที่ใช้บอกความซับซ้อน เช่น ค่าความแตกต่างสัมบูรณ์เฉลี่ย
2. โมเดลนี้ยังได้รวมถึงข้อมูลส่วนหัว (overhead data) และเวกเตอร์การเคลื่อนที่จึงต้องลบออกจากจำนวนบิตที่จะเข้ารหัสเฟรม เพื่อให้ได้จำนวนบิตที่เข้ารหัสเฟรมนั้นอย่างแท้จริง ดังนั้นการหาค่าควอนไทซ์โดยพิจารณา 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 แสดงดังสมการที่ (2.14)

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

เมื่อ MAD_f คือ ค่าความแตกต่างสัมบูรณ์เฉลี่ยของเฟรมปัจจุบัน
 MAD_{f-1} คือ ค่าความแตกต่างสัมบูรณ์เฉลี่ยของเฟรมก่อนหน้า
 a_1, a_2 คือ ค่าสัมประสิทธิ์ โดยเริ่มต้นกำหนดให้มีค่าเท่ากับ 1 และ 0 ตามลำดับ
 H_i คือข้อมูลส่วนหัวและเวกเตอร์การเคลื่อนที่สำหรับเฟรม i

ดังนั้นจากสมการที่ (2.11) สามารถหาคำนวนค่าควอนไทซ์ดังแสดงในสมการที่ (2.15)

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

เมื่อทำการคำนวณหาค่าควอนไทซ์ที่จะทำการเข้ารหัสเฟรมปัจจุบัน (\hat{Q}_{pc}) เพื่อให้คุณภาพสัญญาณวีดิทัศน์ราบเรียบต่อเนื่องกันมากขึ้นจึงทำการปรับค่าควอนไทซ์พารามิเตอร์ไม่ให้ค่าควอนไทซ์ที่จะทำการเข้ารหัสเฟรมปัจจุบันกับค่าควอนไทซ์ที่ทำการเข้ารหัสเฟรมก่อนหน้าต่างกันเกิน 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} คือค่าควอนไทซ์ที่ทำการเข้ารหัสเฟรมก่อนหน้า

โดยค่าควอนไทซ์ที่ได้จะถูกปรับให้ค่าไม่เกินค่าต่ำสุดและค่าสูงสุดของขั้นการควอนไทซ์ดังแสดงในสมการที่ (2.17)

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

เมื่อได้ค่าควอนไทล์แล้วจึงนำไปใช้เพื่อหาโหมด (MODE) ที่เหมาะสมสำหรับแต่ละมาโครบล็อก [23] ดังแสดงในสมการที่ (2.18)

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

เมื่อ $\lambda_{mode} = 0.85 \times 2^{QP/3}$ สำหรับเฟรมแบบ I และ P

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) มาโครบล็อกที่อยู่ติดกันค่าควอนไทล์เปลี่ยนแปลงไม่เกิน c ระดับ

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

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

โดย $c_Y[x, y, MODE | QP]$ คือ ความเข้มแสงของมาโครบล็อกที่ถูกสร้างขึ้น $s_Y[x, y]$ คือ ความเข้มแสงของมาโครบล็อกต้นแบบ c_U c_V s_U และ s_V คือ สัญญาณสีของมาโครบล็อก ส่วน $R(s, c | MODE)$ คือ จำนวนบิตที่ได้จากการเลือกโหมดและค่าควอนไทล์

2.3.2.2 ส่วน Post-Encoding Stage

ทำการปรับค่าพารามิเตอร์ต่างๆ เช่น a_1 และ a_2 ในโมเดลเชิงเส้น b_1 และ b_2 ในโมเดลกำลังสองสำหรับหาค่าควอนไทล์พารามิเตอร์

2.3.3 การควบคุมอัตราระดับหน่วยพื้นฐานย่อย

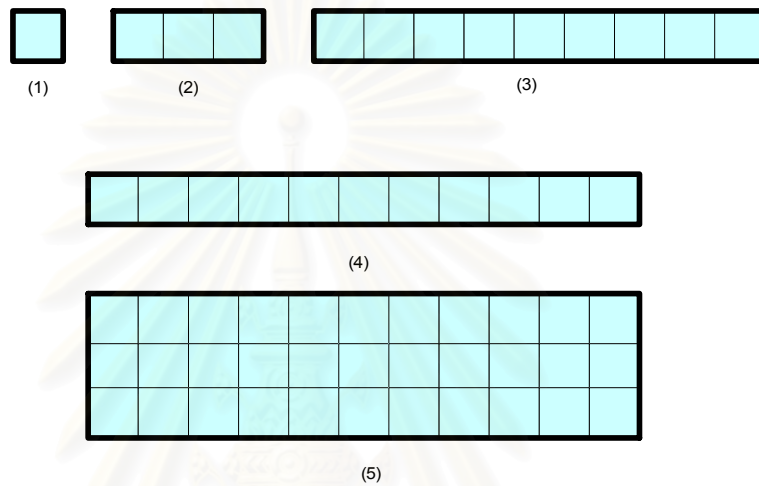
ในตอนนี้จะกล่าวถึงนิยามของหน่วยพื้นฐานย่อยและการควบคุมอัตราระดับหน่วยพื้นฐานย่อย

2.3.3.1 นิยามของหน่วยพื้นฐานย่อย (Basic unit)

ในหนึ่งเฟรมประกอบด้วยมาโครบล็อกหลายๆมาโครบล็อก เช่น ภาพแบบ QCIF จะประกอบ 99 มาโครบล็อก โดยขนาดหนึ่งหน่วยพื้นฐานย่อย (N_{mbpic}) คือ กลุ่มของมาโครบล็อก

ที่มีขนาดตั้งแต่ 1 3 9 11 และ 33 มาโครบล็อกแสดงดังรูปที่ 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

จุดมุ่งหมายของการควบคุมอัตราของหน่วยพื้นฐานย่อย คือ การหาค่าควอนไทล์พารามิเตอร์ที่เหมาะสม โดยที่ผลรวมของจำนวนบิตที่ใช้ในแต่ละเฟรมต้องใกล้เคียงกับจำนวนบิตเป้าหมาย $f(n_{i,j})$ ในแต่ละเฟรมมากที่สุด โดยประกอบด้วย 6 ขั้นตอน ดังนี้

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

ขั้นตอนที่ 1.1 คำนวณจำนวนบิตเป้าหมายสำหรับหน่วยพื้นฐานย่อยปัจจุบันเมื่อ $f_{rb}(n_{i,j,k})$ คือจำนวนบิตคงเหลือสำหรับการเข้ารหัสหน่วยพื้นฐานย่อยตั้งแต่ k ถึง N_{mbunit} และ N_{ub} คือ จำนวนหน่วยพื้นฐานย่อยคงเหลือ 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 $\hat{m}_{hdr,k}$ คือ จำนวนบิตของส่วนหัวแท้จริง สำหรับหน่วยพื้นฐานย่อยที่ k ในเฟรมปัจจุบัน $m_{hdr,1}$ คือ จำนวนบิตของส่วนหัวที่มาจากการประมาณเฟรมก่อนหน้า

ขั้นตอนที่ 1.3 คำนวณจำนวนบิตที่ใช้สำหรับหน่วยพื้นฐานย่อยปัจจุบัน R_{cb} จากสมการที่ (2.22)

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

ขั้นตอนที่ 2 ทำนายค่าความแตกต่างสัมบูรณ์เฉลี่ยสำหรับหน่วยพื้นฐานย่อยปัจจุบันโดยใช้โมเดลเชิงเส้น โดยคำนวณเทียบจากตำแหน่งเดียวกันของเฟรมก่อนหน้าดังสมการ (2.23)

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

เมื่อ MAD_{cb} คือ ค่าความแตกต่างสัมบูรณ์เฉลี่ยของหน่วยพื้นฐานปัจจุบัน MAD_{cb-1} คือ ค่าความแตกต่างสัมบูรณ์เฉลี่ยของหน่วยพื้นฐานก่อนหน้า a_1, a_2 คือ ค่าสัมประสิทธิ์ โดยเริ่มต้นกำหนดให้มีค่าเท่ากับ 1 และ 0 ตามลำดับ

ขั้นตอนที่ 3 คำนวณค่าควอนไทล์พารามิเตอร์โดยใช้โมเดลกำลังสองเหมือนในระดับกลุ่มเฟรม โดยพิจารณาได้ 3 กรณี

กรณีที่ 1 ค่าควอนไทล์พารามิเตอร์สำหรับหน่วยพื้นฐานแรกในเฟรมสามารถหาได้จากสมการที่ (2.24)

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

เมื่อ Q_{apf} คือ ค่าเฉลี่ยค่าควอนไทล์พารามิเตอร์ทั้งเฟรมของเฟรมก่อนหน้า

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

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

เมื่อ Q_{pb} คือ ค่าควอนไทล์พารามิเตอร์ของหน่วยพื้นฐานย่อยก่อนหน้า

$DQuant$ มีค่าเท่ากับ 1

ปรับค่าควอนไทล์พารามิเตอร์เพื่อให้คุณภาพสัญญาณวีดิทัศน์ราบเรียบต่อเนื่องกันมากขึ้นและปรับค่าควอนไทล์ที่ได้ต้องไม่เกินค่าต่ำสุดและค่าสูงสุดของขั้นการควอนไทล์ดังกล่าวที่ (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)

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

ปรับค่าควอนไทล์ที่ได้ต้องไม่เกินค่าต่ำสุดและค่าสูงสุดของขั้นการควอนไทล์ดังสมการที่ (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 คือ ค่าความผิดพลาดกำลังสองเฉลี่ย 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 $M \times N$ คือ ขนาดของเฟรม

$X(m, n)$ คือ ค่าความเข้มแสงของเฟรมปัจจุบันที่จุด (m,n)

$X^R(m, n)$ คือ ค่าความเข้มแสงของเฟรมอ้างอิงที่จุด(m,n)

2.4.2 ค่าอัตราส่วนสัญญาณยอดต่อสัญญาณรบกวน (Peak Signal to Noise Ratio: PSNR)

เป็นค่าที่วัดคุณภาพในแบบ Objective ที่พยายามให้ความหมายในเชิง Subjective [27] 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

$M \times N$ คือ ขนาดของเฟรม

$X(m, n)$ คือ ค่าความเข้มแสงของเฟรมปัจจุบันที่จุด (m,n)

$X^R(m, n)$ คือ ค่าความเข้มแสงของเฟรมอ้างอิงที่จุด(m,n)

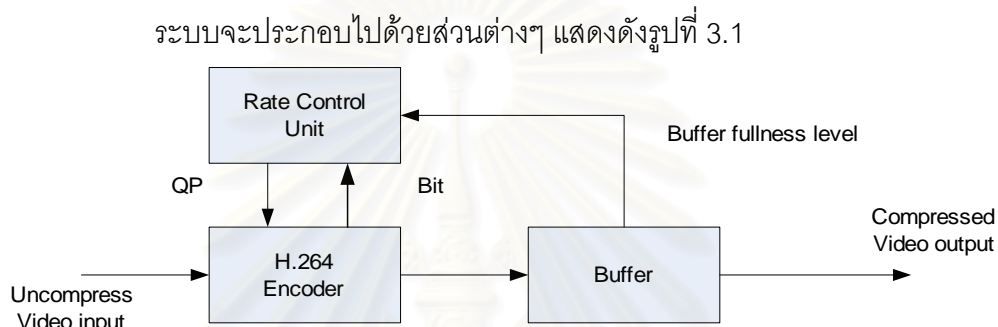
สำหรับวิทยานิพนธ์ฉบับนี้ใช้ค่าอัตราส่วนยอดต่อสัญญาณรบกวน (PSNR) เป็นค่าที่ใช้วัดประสิทธิภาพโดยเรียกว่าค่าคุณภาพสัญญาณวิดีโอที่ค้น

บทที่ 3

รายละเอียดเกี่ยวกับวิทยานิพนธ์และระบบที่นำเสนอ

วิทยานิพนธ์นี้นำเสนอปรับปรุงวิธีการเข้ารหัสสัญญาณวิดีโอที่ต้นตามมาตรฐานการลงรหัสสัญญาณวิดีโอที่ต้นขั้นสูง เพื่อให้สามารถสื่อสารที่อัตราบิตต่ำได้อย่างมีประสิทธิภาพ เมื่อจำกัดการประวิงเวลาไว้ที่ 100 มิลลิวินาที โดยมีรายละเอียดส่วนต่าง ๆ ดังนี้

3.1 ภาพรวมของระบบ



รูปที่ 3.1 ภาพรวมของระบบ

ส่วนแรก คือ ลำดับภาพวิดีโอที่ต้นที่ไม่มีการบีบอัด เมื่อผ่านการเข้ารหัสโดยมาตรฐานการลงรหัสสัญญาณวิดีโอที่ต้นขั้นสูง โดยเมื่อตัวเข้ารหัสกำหนดจำนวนบิตที่ต้องการให้กับหน่วยการควบคุมอัตรา(Rate Control Unit) หน่วยการควบคุมอัตราจะทำการหาค่าควอนไทซ์พารามิเตอร์ที่เหมาะสม โดยจำนวนบิตที่ต้องการสอดคล้องกับจำนวนบิตที่ออกสู่ช่องสัญญาณได้และจำนวนบิตสะสมในบัฟเฟอร์ โดยจำนวนบิตที่ไม่สามารถออกสู่ช่องสัญญาณได้จะถูกเก็บไว้ในบัฟเฟอร์โดย ถ้ามีจำนวนบิตอยู่ในบัฟเฟอร์มาก หน่วยการควบคุมอัตราจะลดจำนวนบิตที่เกิดจากตัวเข้ารหัสเพื่อป้องกันบัฟเฟอร์เต็ม ในกรณีถ้ามีจำนวนบิตอยู่ในบัฟเฟอร์น้อย หน่วยการควบคุมอัตราจะเพิ่มจำนวนบิตที่เกิดจากตัวเข้ารหัสเพื่อป้องกันกรณีไม่มีข้อมูลในบัฟเฟอร์

3.2 ลักษณะสัญญาณวิดีโอที่ต้นขาเข้า

สัญญาณวิดีโอที่ต้นขาเข้าที่ใช้ในวิทยานิพนธ์นี้ คือ ลำดับภาพทดสอบมาตรฐาน (Standard test sequence) ที่อยู่ในรูปแบบ QCIF (176x144 จุดภาพ) ซึ่งมีลักษณะการประยุกต์ใช้งานแบบโทรศัพท์ภาพ และการประชุมสัมมนาทางวิดีโอ ดังรูปที่ 3.2



(ก)



(ข)



(ค)



(ง)

รูปที่ 3.2 สัญญาณวีดิทัศน์ที่ใช้ทดสอบในวิทยานิพนธ์

- (ก) Foreman มีจำนวนเฟรมทั้งหมด 400 เฟรม
- (ข) Silent มีจำนวนเฟรมทั้งหมด 300 เฟรม
- (ค) Carphone มีจำนวนเฟรมทั้งหมด 381 เฟรม
- (ง) Salesman มีจำนวนเฟรมทั้งหมด 400 เฟรม

3.3 การทำงานของระบบที่นำเสนอ

การสื่อสารสัญญาณวีดิทัศน์แบบเวลาจริง สิ่งที่สำคัญคือ การประวิงเวลาควรมีค่าน้อย สื่อสารได้อย่างถูกต้องและต่อเนื่อง เนื่องจากขนาดบัฟเฟอร์มีค่าน้อยถ้าจำนวนบิตที่เข้ารหัสแต่ละเฟรมมีขนาดมากกว่าจำนวนบิตที่คาดหมายไว้จำนวนบิตที่เกินจะถูกเก็บสะสมไว้ในบัฟเฟอร์ของตัวเข้ารหัส และเพิ่มการประวิงเวลาของบัฟเฟอร์ (buffer delay) ถ้าข้อมูลสะสมในบัฟเฟอร์มาก จะทำให้บัฟเฟอร์มากเกินเก็บ (buffer overflow) โดยปกติเพื่อลดการประวิงเวลาและจำนวนบิตสะสมในบัฟเฟอร์ ตัวเข้ารหัสจะทำการกระโดดข้ามเฟรม (frame skipping) ซึ่งเป็นสาเหตุของความไม่ต่อเนื่องด้านการเคลื่อนไหวของสัญญาณวีดิทัศน์ ดังนั้นวิทยานิพนธ์นี้นำเสนอการควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูง (AVC) เพื่อสามารถสื่อสารที่อัตราบิตต่ำได้อย่างมีประสิทธิภาพ โดยได้ศึกษาผลกระทบของการสะสมบัฟเฟอร์ พบว่ามีจำนวน

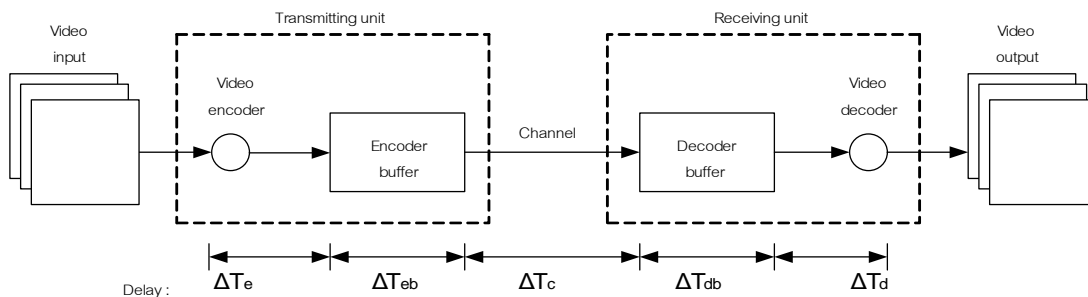
เฟรมกระโดดมากและลดค่าคุณภาพสัญญาณวีดิทัศน์ ทั้งนี้เนื่องจากจำนวนบิตเป้าหมายกับจำนวนบิตที่เข้ารหัสจริงแตกต่างกันมาก จึงได้มีการศึกษาพารามิเตอร์ที่มีผลต่อการกำหนดบิตเป้าหมาย ซึ่งพารามิเตอร์ดังกล่าวคือ ค่าถ่วงน้ำหนักบิตเป้าหมาย (β) โดยได้ทำการปรับค่าถ่วงน้ำหนัก พบว่า ค่าถ่วงน้ำหนักบิตเป้าหมายมีผลต่อจำนวนเฟรมกระโดดและค่าคุณภาพสัญญาณวีดิทัศน์ เมื่อปรับค่าให้มีค่ามากขึ้นจำนวนเฟรมกระโดดลดลงและค่าคุณภาพสัญญาณวีดิทัศน์ดีขึ้น ดังนั้น การเลือกค่าถ่วงน้ำหนักที่เหมาะสมเป็นสิ่งสำคัญซึ่งค่าถ่วงน้ำหนักบิตเป้าหมายเปลี่ยนแปลงตามลักษณะภาพวีดิทัศน์จึงได้นำเสนอการปรับค่าถ่วงน้ำหนักแบบปรับตัวได้ โดยจะทำการปรับค่าถ่วงน้ำหนักบิตเป้าหมายทุกเฟรมที่เข้ารหัส

อย่างไรก็ตาม การปรับค่าถ่วงน้ำหนักบิตเป้าหมาย สามารถลดจำนวนเฟรมกระโดดส่วนหนึ่งเท่านั้นเนื่องจากการกำหนดค่าควอนไทซ์พารามิเตอร์มิได้คำนึงถึงจำนวนบิตสะสมในบัพเฟอร์ในขณะนั้น ดังนั้น จึงนำเสนอเงื่อนไขสำหรับการกำหนดค่าควอนไทซ์พารามิเตอร์โดยคำนึงถึงจำนวนบิตสะสมในบัพเฟอร์ในขณะนั้นเพื่อป้องกันมิให้มีจำนวนบิตสะสมในบัพเฟอร์มากกว่าค่าสูงสุดบัพเฟอร์ โดยทดสอบร่วมกับการกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายแบบคงที่และแบบปรับตัวได้ โดยผลการทดสอบพบว่า การปรับค่าถ่วงน้ำหนักแบบปรับตัวได้สามารถให้ประสิทธิภาพดีกว่า

3.3.1 การศึกษาผลกระทบของการสะสมบัพเฟอร์ของการควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูงต่อคุณภาพสัญญาณวีดิทัศน์เมื่อจำกัดการประวิงเวลา

ในส่วนนี้แสดงผลกระทบของการสะสมของบัพเฟอร์ของการควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูงต่อคุณภาพสัญญาณวีดิทัศน์เมื่อจำกัดให้การประวิงเวลา โดยทั่วไปปัจจัยที่มีผลต่อการประวิงเวลาในระบบสื่อสาร [24] แสดงดังรูปที่ 3.3 ประกอบด้วย

1. การประวิงเวลาที่เกิดจากตัวเข้ารหัส ΔT_e (การประวิงเวลาที่เกิดจากการประมวลผล)
2. การประวิงเวลาที่เกิดจากบัพเฟอร์ของตัวเข้ารหัส ΔT_{eb} (การควบคุมอัตราให้คงที่ก่อนส่งออกช่องสัญญาณ)



รูปที่ 3.3 ปัจจัยต่างๆที่มีผลต่อการประวิงเวลาในระบบสื่อสารทางวิดีโอ [24]

3. การประวิงเวลาที่เกิดจากการส่งผ่านช่องสัญญาณ ΔT_c (อาจเกิดจากต้องรอใช้ช่องสัญญาณหรือบางครั้งเกิดจากความผิดพลาดจากการส่งข้อมูลทำให้ต้องส่งรื้อขอให้ส่งข้อมูลใหม่ (retransmission))
4. การประวิงเวลาที่เกิดจากบัฟเฟอร์ของตัวถอดรหัส T_{db} (เกิดจากการควบคุมอัตราให้คงที่ สามารถถอดรหัสได้อย่างไม่สะดุด)
5. การประวิงเวลาที่เกิดจากการประมวลผลของตัวถอดรหัส ΔT_d

ITU-T ได้กำหนดค่าสูงสุดของการประวิงเวลาที่สามารถยอมรับได้ของการสนทนาผ่านโทรศัพท์สากลไว้ในมาตรฐาน G.114 ไม่เกิน 150 มิลลิวินาที [24] ดังนั้นในการจำลองระบบจึงจำกัดค่าการประวิงเวลาไว้ที่ 100 มิลลิวินาทีสำหรับการประวิงเวลาที่เกิดขึ้นเกิดจากการสะสมของบิตในบัฟเฟอร์ของตัวเข้ารหัส

เพื่อแสดงผลกระทบของการสะสมของบัฟเฟอร์ของการควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวิดีโอวีดิทัศน์ขั้นสูงเมื่อจำกัดค่าการประวิงเวลาที่ 100 มิลลิวินาทีต่อคุณภาพสัญญาณวีดิทัศน์ จึงทำการเข้ารหัสสัญญาณวีดิทัศน์โดยใช้ลำดับภาพต่าง ๆ เข้ารหัสที่ 32 กิโลบิตต่อวินาที [25]

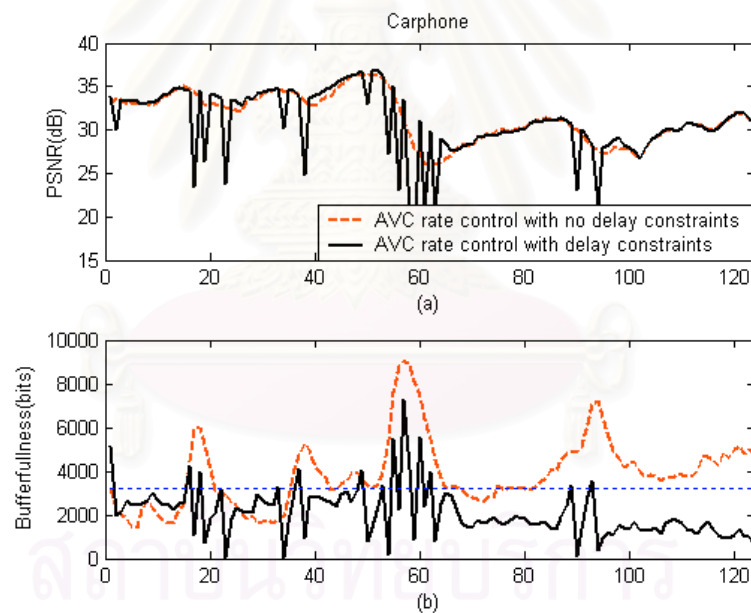
สำหรับการพิจารณาการกระโดดของเฟรมภาพพิจารณาจากจำนวนบิตในบัฟเฟอร์ของเฟรมปัจจุบัน $B_c(n_{i,j})$ ถ้ามีขนาดมากกว่าค่าสูงสุดของขนาดบัฟเฟอร์ ตัวเข้ารหัสจะทำการกระโดดข้ามเฟรมถัดไป เพื่อให้จำนวนบิตในบัฟเฟอร์ที่ค้างอยู่ได้ถูกส่งออกสู่ช่องสัญญาณค่าการประวิงเวลาที่เกิดจากบัฟเฟอร์ตัวเข้ารหัสควรน้อยกว่า $\frac{M}{u(n_{i,j})}$ วินาที โดยทั่วไปจะกำหนดให้

ค่าสูงสุดของขนาดบัฟเฟอร์เท่ากับ $M = \frac{u(n_{i,j})}{F_r}$ [5] และค่าการประวิงเวลาสูงสุด (τ) สามารถคำนวณได้จากสมการที่ (3.1)

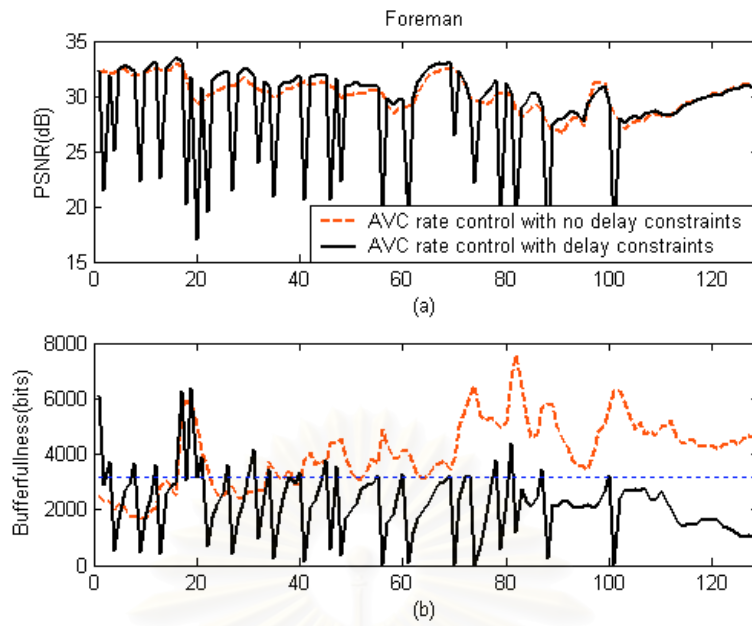
$$\tau = \frac{M}{u(n_{i,j})} = \frac{1}{F_r} \text{วินาที} \quad (3.1)$$

ตัวอย่างเช่น อัตราบิตเท่ากับ 32 กิโลบิตต่อวินาที อัตราเฟรมเท่ากับ 10 เฟรมต่อวินาที ดังนั้นค่าการประวิงเวลาสูงสุดคือ 100 มิลลิวินาที

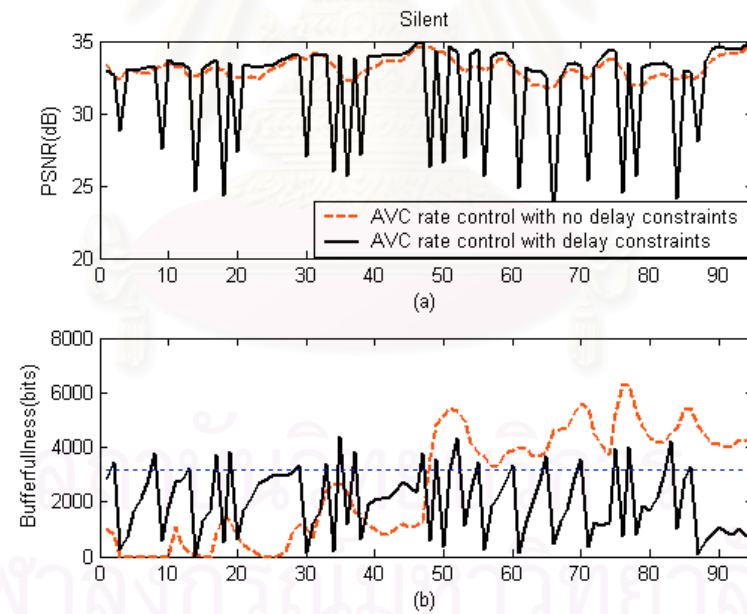
ในกรณีที่เกิดเฟรมกระโดด การคำนวณค่าคุณภาพสัญญาณวีดิทัศน์คำนวณจากนำเฟรมก่อนหน้ามาคำนวณหาค่าคุณภาพสัญญาณวีดิทัศน์แทนเฟรมที่ถูกกระโดดข้ามไปเพราะที่ด้านตัวถอดรหัสยังคงแสดงภาพเดิมอยู่ โดยใช้หลักการคำนวณค่าคุณภาพสัญญาณวีดิทัศน์ตามที่มีกำหนดไว้ในการควบคุมอัตรา MPEG-4 [1]



รูปที่ 3.4 ผลการทดสอบการเข้ารหัสสัญญาณวีดิทัศน์ของลำดับภาพ Carphone (a) ค่าคุณภาพสัญญาณวีดิทัศน์ของควบคุมอัตราสำหรับการลงรหัสสัญญาณวีดิทัศน์ขั้นสูง เมื่อจำกัดการประวิงเวลา (เส้นทึบ) และไม่จำกัดการประวิงเวลา (เส้นประ) และ (b) จำนวนบิตสะสมในบัฟเฟอร์สำหรับแต่ละเฟรมภาพ ค่าสูงสุดของขนาดบัฟเฟอร์ (เส้นประ)



รูปที่ 3.5 ผลการทดสอบการเข้ารหัสสัญญาณวิดีโอที่ศน์ของลำดับภาพ Foreman (a) ค่าคุณภาพสัญญาณวิดีโอที่ศน์ของควบคุมอัตราสำหรับการลงรหัสสัญญาณวิดีโอที่ศน์ขึ้นสูง เมื่อจำกัดการประวิงเวลา (เส้นทึบ) และ ไม่จำกัดการประวิงเวลา (เส้นประ) และ (b) จำนวนบิตสะสมในบัฟเฟอร์สำหรับแต่ละเฟรมภาพ ค่าสูงสุดของขนาดบัฟเฟอร์ (เส้นประ)



รูปที่ 3.6 ผลการทดสอบการเข้ารหัสสัญญาณวิดีโอที่ศน์ของลำดับภาพ Silent (a) ค่าคุณภาพสัญญาณวิดีโอที่ศน์ของควบคุมอัตราสำหรับการลงรหัสสัญญาณวิดีโอที่ศน์ขึ้นสูง เมื่อจำกัดการประวิงเวลา (เส้นทึบ) และ ไม่จำกัดการประวิงเวลา (เส้นประ) และ (b) จำนวนบิตสะสมในบัฟเฟอร์สำหรับแต่ละเฟรมภาพ ค่าสูงสุดของขนาดบัฟเฟอร์ (เส้นประ)

จากรูปที่ 3.4 (a) 3.5 (a) และ 3.6 (a) แสดงค่าคุณภาพสัญญาณวีดิทัศน์ และ รูปที่ 3.4 (b) 3.5 (b) และ 3.6 (b) แสดงจำนวนบิตสะสมในบัพเฟอริในเฟรมต่างๆ พบว่า จำนวนบิตสะสมในบัพเฟอริมีค่ามากกว่าค่าสูงสุดของขนาดบัพเฟอริทั้งในกรณีที่จำกัดการประวิงเวลาและไม่จำกัดการประวิงเวลา แต่ในกรณีไม่จำกัดการประวิงเวลาจะไม่มี การตรวจสอบการกระโดดข้ามเฟรม ในกรณีที่จำกัดการประวิงเวลาต้องมีการตรวจสอบการกระโดดข้ามเฟรมทำให้มีหลายเฟรมภาพไม่สามารถแสดงได้ที่ด้านตัวถอดรหัส เพราะยังมีบิตส่วนหนึ่งยังค้างอยู่ใน บัพเฟอริตัวเข้ารหัส จึงต้องกระโดดข้ามเฟรมเพื่อลดจำนวนบิตที่ค้างอยู่ที่บัพเฟอริของตัวเข้ารหัส ทำให้สัญญาณวีดิทัศน์มีความไม่ต่อเนื่องส่งผลให้คุณภาพสัญญาณวีดิทัศน์ลดลง โดยจำนวนเฟรมที่กระโดดและค่าคุณภาพสัญญาณวีดิทัศน์ลดลงอยู่ในช่วง 0.21 dB ถึง 1.24 dB สามารถแสดงดังตารางที่ 3.1

จากที่แสดงให้เห็นถึงผลกระทบของการสะสมของบัพเฟอริของการควบคุมอัตรา สำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูงต่อคุณภาพสัญญาณวีดิทัศน์เมื่อจำกัดการประวิงเวลา พบว่ามีจำนวนเฟรมกระโดดมากอยู่ในช่วงระหว่าง 15 เฟรม ถึง 21 เฟรม ทำให้การสื่อสารขาดความต่อเนื่อง สามารถแสดงดังรูปที่ 3.7 โดยรูปที่ 3.7(ก) ถึง (ค) เป็นการเข้ารหัสที่ใช้การควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูง และรูปที่ 3.7 (ง) ถึง (จ) เป็นการเข้ารหัสที่ใช้การควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูงเมื่อจำกัดการประวิงเวลา สังเกตได้ว่า รูปที่ 3.7 (จ) ตัวถอดรหัสต้องแสดงภาพเดิมซ้ำเนื่องจากเกิดเฟรมกระโดดที่เฟรม 52

ในวิทยานิพนธ์ฉบับนี้จึงนำเสนอวิธีปรับปรุงคุณภาพสัญญาณวีดิทัศน์ โดยปรับค่าพารามิเตอร์ต่างๆ เพื่อสามารถควบคุมจำนวนบิตที่อยู่ในบัพเฟอริ การควบคุมอัตราที่นำเสนอในระดับกลุ่มภาพ เฟรม และหน่วยพื้นฐานย่อย โดยมีหลักการในแต่ละส่วนดังต่อไปนี้

3.3.2 ส่วนควบคุมการกระโดดของเฟรม

ก่อนที่จะทำการเข้ารหัสเฟรมปัจจุบันจะต้องมีการคำนวณจำนวนบิตสะสมในบัพเฟอริของตัวเข้ารหัส ($B_c(n_i, j)$) ดังสมการที่ (3.2)

$$B_c(n_i, j) = B_c(n_i, j-1) + (A(n_i, j-1) - (\frac{u(n_i, j-1)}{F_r})) \quad (3.2)$$

โดยกำหนดให้ขนาดบัพเฟอร์เริ่มต้นมีค่าเท่ากับ 0 โดยที่การปรับจำนวนบิต
 สะสมในบัพเฟอร์เท่ากับจำนวนบิตสะสมในบัพเฟอร์ของเฟรมก่อนหน้าบวกกับจำนวนบิตที่ใช้
 เข้ารหัสเฟรมก่อนหน้าลบด้วยจำนวนบิตที่ออกสู่ช่องสัญญาณ (ในที่นี้เท่ากับ 3200 บิต) ถ้าจำนวน
 บิตในบัพเฟอร์มีขนาดมากกว่าหรือเท่ากับค่าสูงสุดของบัพเฟอร์ ตัวเข้ารหัสจะทำการกระโดดข้าม
 เฟรมจนกระทั่งจำนวนบิตในบัพเฟอร์มีค่าน้อยกว่าค่าสูงสุดของบัพเฟอร์ ดังสมการที่ (3.3) ถ้ามี
 จำนวนบิตสะสมในบัพเฟอร์มากยิ่งขึ้นทำให้มีการเฟรมกระโดดมากขึ้นเท่านั้น

$$\text{while } B_c(n_{i,j}) > M \quad (3.3)$$

$$\{ B_c(n_{i,j}) = B_c(n_{i,j}) - \frac{u(n_{i,j-1})}{F_r} \}$$

ตารางที่ 3.1 ค่าคุณภาพสัญญาณวีดิทัศน์และจำนวนเฟรมกระโดดของการควบคุมอัตรา
 สำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูงและ การควบคุมอัตราสำหรับมาตรฐานการ
 ลงรหัสสัญญาณวีดิทัศน์ขั้นสูงเมื่อจำกัดการประวิงเวลา

ลำดับภาพ	การควบคุมอัตราสำหรับ มาตรฐานการลงรหัส สัญญาณวีดิทัศน์ขั้นสูง	การควบคุมอัตราสำหรับมาตรฐานการลงรหัส สัญญาณวีดิทัศน์ขั้นสูง เมื่อจำกัดการประวิงเวลา	
	ค่าคุณภาพสัญญาณ วีดิทัศน์ (dB)	จำนวนเฟรมกระโดด	ค่าคุณภาพสัญญาณ วีดิทัศน์ (dB)
Carphone	31.52	15	30.90(-0.62)
Forman	30.28	21	29.04(-1.24)
Silent	33.14	20	32.09(-1.05)
Salesman	34.78	16	34.57(-0.21)



(ก) เฟอร์มที่ 51



(ข) เฟอร์มที่ 52



(ค) เฟอร์มที่ 53



(ง) เฟอร์มที่ 51



(จ) เฟอร์มที่ 52



(ฉ) เฟอร์มที่ 53

รูปที่ 3.7 เปรียบเทียบการเข้ารหัสลำดับภาพ Foreman (ก) ถึง (ค) ใช้การควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูงและ (ง) ถึง (จ) ใช้การควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูง เมื่อจำกัดการประวิงเวลา

3.3.3 ส่วนการควบคุมอัตราในระดับกลุ่มภาพ

สำหรับชนิดของภาพ ในวิทยานิพนธ์นี้ลำดับภาพแบบ IPPP ซึ่งลำดับภาพแบบนี้ใช้สำหรับการเข้ารหัสเมื่อจำกัดการประวิงเวลา

ในระดับนี้คำนวณหาจำนวนบิตที่เหลือสำหรับเข้ารหัสเฟรมที่เหลือในแต่ละกลุ่มภาพ (T_r) เหมือนกับในระดับกลุ่มภาพของการควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูง ดังแสดงในสมการที่ (3.4)

$$T_r(n_{1,0}) = \frac{u(n_{1,1})}{F_r} * N_{gop} \quad (3.4)$$

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

T_r คือ จำนวนบิตสำหรับกลุ่มภาพ

N_{gop} คือ จำนวนเฟรมในแต่ละกลุ่มภาพ (GOP)

F_r คือ อัตราเฟรม

$B_c(n_{1,j})$ คือจำนวนบิตที่อยู่ในบัฟเฟอร์หลังจากเข้ารหัสแต่ละเฟรม กำหนดให้ขนาดบัฟเฟอร์เริ่มต้นมีค่าเท่ากับ 0

$u(n_{1,j})$ คือ อัตราในช่องสัญญาณ ในเฟรมที่ j ในกลุ่มภาพที่ 1

โดยการคำนวณหาจำนวนบิตสะสมในบัฟเฟอร์สามารถคำนวณโดยใช้โมเดล ทราฟฟิกการไหลของไหล โดยจำนวนบิตในบัฟเฟอร์หลังจากเข้ารหัสเฟรม j ($B_c(n_{1,j+1})$) จะเลือกค่ามากที่สุดระหว่างค่าศูนย์กับจำนวนบิตในบัฟเฟอร์ของเฟรมก่อนหน้ากับผลต่างระหว่างจำนวนบิตที่ใช้ในการเข้ารหัสเฟรมก่อนหน้ากับจำนวนบิตที่สามารถออกสู่ช่องสัญญาณได้ ดังแสดงในสมการที่ (3.5)

$$B_c(n_{1,j+1}) = \max\{0, B_c(n_{1,j}) + A(n_{1,j}) - \frac{u(n_{1,j})}{F_r}\} \quad (3.5)$$

โดยที่ $A(n_{1,j})$ คือ จำนวนบิตที่ใช้ในการเข้ารหัสเฟรมที่ j ในกลุ่มภาพที่ 1

ดังนั้นถ้าทำการเข้ารหัสแต่ละเฟรมเสร็จต้องทำการคำนวณจำนวนบิตคงเหลือที่สามารถคำนวณดังแสดงในสมการที่ (3.6)

$$T_r(n_{1,j}) = T_r(n_{1,j-1}) - A(n_{1,j-1}) \quad (3.6)$$

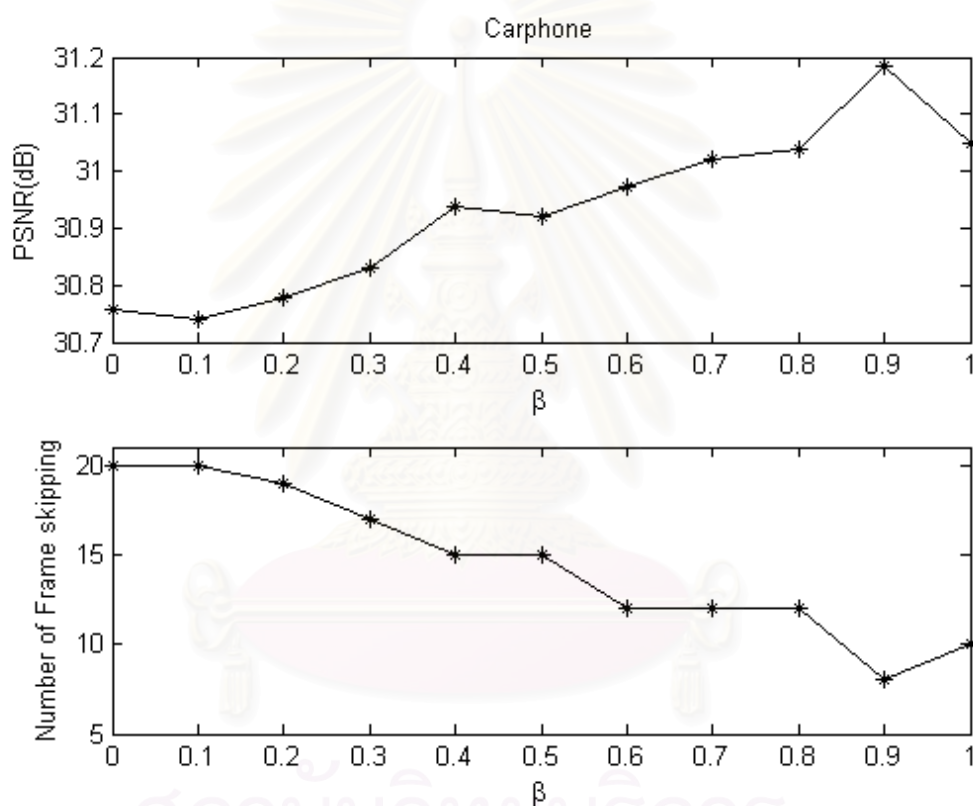
3.3.4 ส่วนการควบคุมอัตราในระดับเฟรม

ในระดับนี้แบ่งได้ 2 ขั้นตอน

3.3.4.1 ส่วน Pre-Encoding Stage

จากที่กล่าวข้างต้นในส่วนการควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณ วิดีทัศน์ขั้นสูง จำนวนบิตเป้าหมายสามารถคำนวณได้จาก 2 ส่วน คือ ส่วนที่ใช้ระดับบัฟเฟอร์ และ ส่วนที่ใช้จำนวนบิตคงเหลือจากระดับกลุ่มภาพ แล้วถ่วงน้ำหนักด้วยค่าถ่วงน้ำหนักบิตเป้าหมาย ในวิทยานิพนธ์นี้เสนอให้มีการปรับค่าถ่วงน้ำหนักบิตเป้าหมาย โดยปรับให้มีค่าตั้งแต่ 0 ถึง 1 โดย

เพิ่มทีละ 0.1 สำหรับการกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายเท่ากับศูนย์ เปรียบได้เท่ากับ การกำหนดบิตเป้าหมาย โดยใช้จำนวนบิตคงเหลือจากระดับกลุ่มภาพเพียงส่วนเดียว และค่าถ่วงน้ำหนักบิตเป้าหมาย เท่ากับหนึ่ง เปรียบได้เหมือนกับการกำหนดบิตเป้าหมายโดยใช้ส่วนที่ค้ำึงถึงระดับบัพเฟอร์เพียงส่วนเดียว อย่างไรก็ตาม การเลือกค่าถ่วงน้ำหนักบิตเป้าหมายที่เหมาะสมเป็นสิ่งสำคัญ เมื่อเราปรับค่าถ่วงน้ำหนักบิตเป้าหมาย ซึ่งค่าถ่วงน้ำหนักบิตเป้าหมาย ที่แตกต่างกันมีผลต่อจำนวนเฟรมกระโดดและค่าคุณภาพสัญญาณวีดิทัศน์ในลำดับภาพ Carphone ผลที่ได้แสดงดังรูปที่ 3.8 ดังนั้นการกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายควรปรับตามกิจกรรมการเคลื่อนที่ (Motion Activity) และจำนวนบิตสะสมในบัพเฟอร์



รูปที่ 3.8 ค่าคุณภาพสัญญาณวีดิทัศน์และจำนวนเฟรมกระโดดเมื่อปรับค่าถ่วงน้ำหนักบิตเป้าหมายต่าง ๆ

ในวิทยานิพนธ์นี้นำเสนอ การปรับตัวของค่าถ่วงน้ำหนักบิตเป้าหมาย เพื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายให้เหมาะสมกับลำดับภาพต่าง ๆ โดยวิธีที่นำเสนอจะปรับค่าถ่วงน้ำหนักบิตเป้าหมายทุกเฟรมที่เข้ารหัส ขั้นตอนการปรับค่ามีดังนี้

ขั้นตอนที่ 1 ทำการหากิจกรรมการเคลื่อนที่ (motion activity) [26] โดยหาได้จากผลรวมของค่าเวกเตอร์การเคลื่อนที่ในแนวแกน X และแกน Y โดยหาได้จากสมการที่ (3.7)

$$MA = \sum_{i=1}^N |(MV_i(X))| + |(MV_i(Y))| \quad (3.7)$$

เมื่อ N คือ จำนวนมาโครบล็อกในเฟรม

ขั้นตอนที่ 2 หาค่าถ่วงน้ำหนักบิตเป้าหมายที่เหมาะสมโดยมีแนวความคิดในการตัดสินใจ คือ ถ้าค่าการเปลี่ยนแปลงการเคลื่อนที่เฟรมปัจจุบันมีค่าสูงกว่าค่าการเปลี่ยนแปลงการเคลื่อนที่ของเฟรมก่อนหน้ามาก ๆ และจำนวนบิตสะสมในบัพเฟอร์สูงกว่า 75% ของค่าสูงสุดบัพเฟอร์ จะกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายต่ำกว่าค่า $\beta_{initial}$ และถ้าค่าการเปลี่ยนแปลงการเคลื่อนที่เฟรมปัจจุบันมีค่าต่ำกว่าค่าการเปลี่ยนแปลงการเคลื่อนที่ของเฟรมก่อนหน้ามาก ๆ และจำนวนบิตสะสมในบัพเฟอร์น้อยกว่า 75% ของค่าสูงสุดบัพเฟอร์ จะกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายต่ำกว่าค่า $\beta_{initial}$ ซึ่งมีกระบวนการดังนี้

$$\text{if}((MA - PMA) > \text{Threshold} \quad \text{and} \quad (B_c(n_{1,j})) > 75\%M)$$

$$\beta = \beta_{initial} + 0.1$$

$$\text{else if}((MA - PMA < -\text{Threshold}) \quad \text{and} \quad (B_c(n_{1,j})) < 75\%M)$$

$$\beta = \beta_{initial} - 0.1$$

else

$$\beta = \beta_{initial}$$

3.3.4.2 ส่วน Post-Encoding Stage

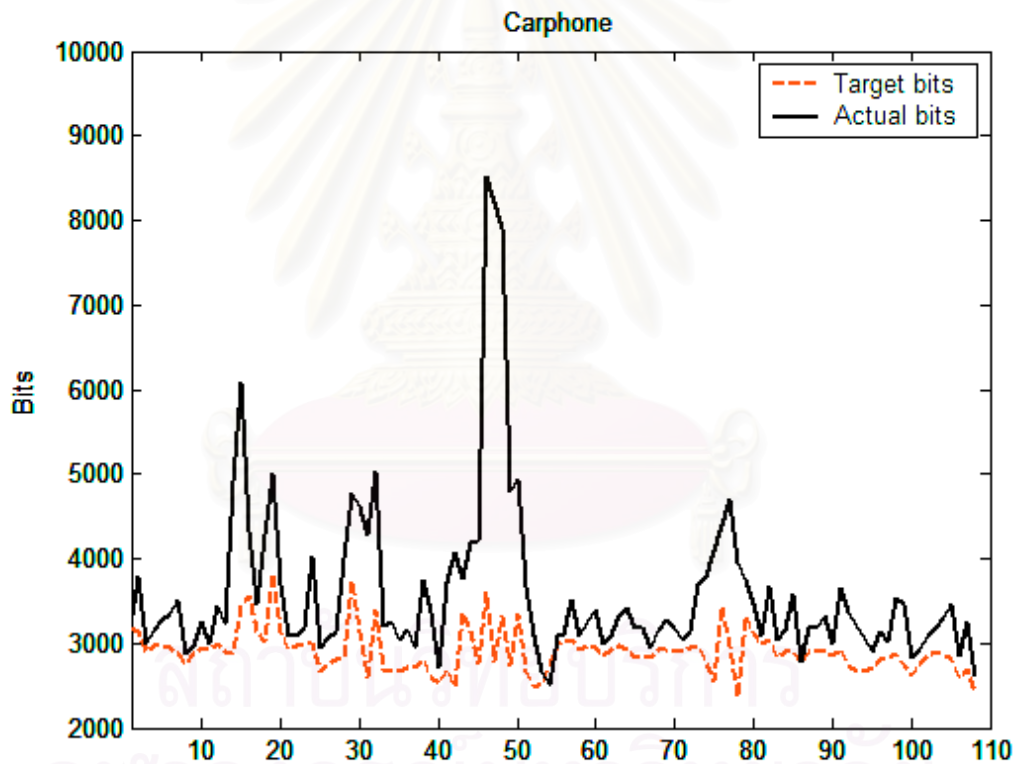
ตรวจสอบจำนวนบิตที่อยู่ในบัพเฟอร์ว่าต้องทำกระโดดข้ามเฟรมหรือไม่ในส่วนของควบคุมการกระโดดของเฟรมถ้ามีขนาดมากกว่าขนาดบัพเฟอร์ตัวเข้ารหัสจะทำการกระโดดข้ามเฟรมถัดไปที่จะทำการเข้ารหัส เพื่อให้จำนวนบิตในบัพเฟอร์ที่ค้างอยู่ได้ออกสู่ของสัญญาณจนกว่ามีค่าน้อยกว่าค่าสูงสุดของบัพเฟอร์เมื่อใช้เงื่อนไขป้องกันการล้นของบัพเฟอร์ทำให้จำนวนเฟรมกระโดดน้อยลง

3.3.4 การควบคุมอัตราในระดับหน่วยพื้นฐานย่อย

สำหรับการเลือกขนาดหน่วยพื้นฐานย่อย เลือกขนาดหน่วยพื้นฐานย่อยเท่ากับ 1 (1 เฟรม มีทั้งหมด 99 หน่วยพื้นฐานย่อย) เพราะขนาดหน่วยพื้นฐานย่อยที่เล็กสามารถควบคุมจำนวนบิตที่ใช้ในการเข้ารหัสในระดับที่ละเอียดกว่า

ในวิทยานิพนธ์นี้ จะใช้วิธีการควบคุมอัตราที่หน่วยพื้นฐานย่อยของมาตรฐานการเข้ารหัสสัญญาณวีดิทัศน์ขั้นสูง ที่อธิบายในหัวข้อที่ 2.3.3.2 โดยนำเสนอขั้นตอนที่ 3 กรณีที่ 2 โดยมีรายละเอียดดังนี้

กรณีที่ 2 ถ้าจำนวนบิตเป้าหมาย (f_{rb}) มีค่าน้อยกว่า 0 ทำการเพิ่มค่าระดับควอนไทซ์พารามิเตอร์เพื่อไม่ให้ผลรวมจำนวนบิตที่เข้ารหัสเกินกว่าจำนวนบิตเป้าหมายมาก แต่ค่าควอนไทซ์พารามิเตอร์ถูกจำกัดไม่ให้เกินกว่า 3 ระดับ ซึ่งทำให้ผลรวมจำนวนบิตเกินกว่าจำนวนบิตเป้าหมายมากดังแสดงดังรูปที่ 3.9



รูปที่ 3.9 จำนวนบิตเป้าหมาย $f(n_{i,j})$ และจำนวนบิตที่เข้ารหัสจริง $A(n_{i,j})$

สำหรับลำดับภาพ Carphone

จึงได้มีการกำหนดเงื่อนไขการเลือกค่าควอนไทซ์พารามิเตอร์ โดยคำนึงถึงจำนวนบิตในบัฟเฟอร์ขณะนั้น (B_k) โดยหลังจากเข้ารหัสหน่วยพื้นฐานย่อย k ดังสมการที่ (3.8)

$$B_k = B_c(n_{1,j}) + \sum_{i=1}^k B_{b-i}, 0 \leq B_k \leq 2M \quad (3.8)$$

เมื่อ B_{b-i} คือ จำนวนบิตที่เข้ารหัสตั้งแต่หน่วยพื้นฐานย่อยที่ 1 ถึง k

โดยถ้าจำนวนบิตในบัฟเฟอร์ขณะนั้นมีค่าน้อยกว่าระดับจุดเริ่มเปลี่ยนบัฟเฟอร์ (Buffer Threshold: h) ดังนั้น ค่าควอนไทซ์พารามิเตอร์จะมีค่าเท่ากับสมการที่ (2.26) แต่ถ้ามากกว่า ค่าควอนไทซ์พารามิเตอร์จะใช้ค่าเดิม

if ($f(n_{1,j}) < 0$)

$$\hat{Q}_{cb} = Q_{pb} + DQUANT$$

if ($B_k < hM$) {

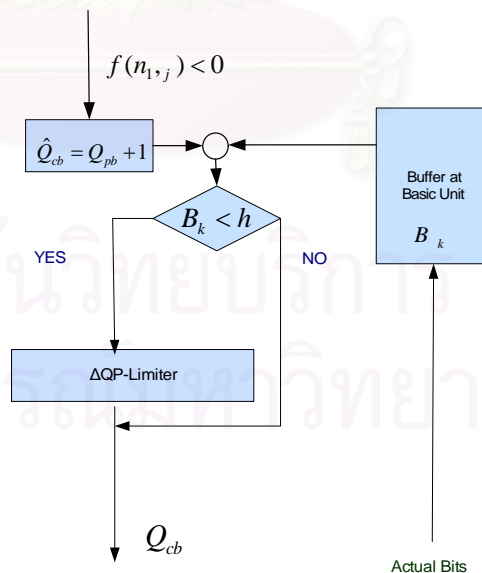
$$Q_{cb} = \max\{1, Q_{apf} - \Delta, \min\{51, Q_{apf} + \Delta, \hat{Q}_{cb}\}\}$$

}

else

$$Q_{cb} = \hat{Q}_{cb}$$

สามารถแสดงดังรูปที่ 3.10



รูปที่ 3.10 การควบคุมอัตราในระดับหน่วยพื้นฐานย่อยที่นำเสนอ

บทที่ 4

ผลการทดสอบ

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

4.1 ขั้นตอนการทดสอบ

การควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูง มีการทำงานดังรูปที่ 2.10 โดยเมื่อจำกัดการประวิงเวลาที่ 100 มิลลิวินาทีแล้วพบว่ามีความเฟรมกระโดดมากทำให้ค่าคุณภาพสัญญาณวีดิทัศน์ลดลง ดังนั้นวิทยานิพนธ์นี้จึงนำเสนอการปรับอัตราบิตให้แก่สัญญาณวีดิทัศน์โดยปรับค่าให้สอดคล้องกับจำนวนบิตสะสมในบัฟเฟอร์เพื่อคงคุณภาพที่ดีของสัญญาณ ทั้งนี้วิธีที่นำเสนออยู่บนพื้นฐานการมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูง

4.1.1 การกำหนดพารามิเตอร์ในการเข้ารหัส

1. การกำหนดอัตราเฟรมในการเข้ารหัส ในการทดลองนี้กำหนดใช้อัตราเฟรมเท่ากับ 10 เฟรมต่อวินาที ซึ่งเป็นอัตราเฟรมที่ใช้ในการเข้ารหัสสัญญาณวีดิทัศน์ที่อัตราบิตต่ำ
2. การกำหนดชนิดของสัญญาณวีดิทัศน์ที่ใช้ในการเข้ารหัส ซึ่งในที่นี้เป็น QCIF มีขนาด 176x144 จุดภาพ
3. การเลือกชนิดของภาพ ใช้เฟรมแรกเป็นเฟรม I และที่เหลือใช้เฟรมแบบ P ตลอดการเข้ารหัส
4. กำหนดอัตราช่องสัญญาณในการส่งที่ 32 กิโลบิตต่อวินาที
5. ค่าควอนไทซ์สำหรับเฟรม I เฟรมแรกเท่ากับ 32

ค่าพารามิเตอร์อื่น ๆ ที่ใช้ในการทดลองแสดงดังตารางที่ 1

ตารางที่ 4.1 พารามิเตอร์ที่ใช้ในการทดสอบ

MV resolution	1/4 pel
Hadamard	on
RD optimization	off
Search Range	± 16
Reference frames	1
Restrict Search Range	2
Symbol mode	UVLC
GOP structure	IPPP

4.2.3.1 ลำดับภาพที่ใช้ในการทดสอบ

โดยลักษณะโดยทั่วไปของชุดภาพต่าง ๆ ที่ใช้ทดสอบ มีดังนี้

1. ลำดับภาพ Carphone มีการเคลื่อนไหวของวัตถุเร็ว มีการเปลี่ยนแปลงของพื้นหลัง และเคลื่อนไหวที่ของคน
ค่ากิจกรรมการเคลื่อนไหวที่อยู่ระหว่าง: 158 ถึง 2518
2. ลำดับภาพ Foreman มีการเคลื่อนไหวที่ของวัตถุด้วยความเร็วปานกลาง มีการเคลื่อนไหวที่และสาย (pan) กว้าง
ค่ากิจกรรมการเคลื่อนไหวที่อยู่ระหว่าง: 143 ถึง 13090
3. ลำดับภาพ Silent มีการเคลื่อนไหวที่ของวัตถุด้วยความเร็ว การเคลื่อนไหวที่ส่วนใหญ่อยู่ที่มือและใบหน้า
ค่ากิจกรรมการเคลื่อนไหวที่อยู่ระหว่าง: 46 ถึง 2936
4. ลำดับภาพ Salesman มีการเคลื่อนไหวของวัตถุช้า การเคลื่อนไหวที่ส่วนใหญ่อยู่บริเวณมือและใบหน้า
ค่ากิจกรรมการเคลื่อนไหวที่อยู่ระหว่าง: 5 ถึง 423

4.2.1.1 ผลการทดสอบและวิเคราะห์ผลการเข้ารหัส

เพื่อพิจารณาระบบในหลาย ๆ แง่มุม ไม่ว่าจะเป็นคุณภาพของสัญญาณหรือประสิทธิภาพในการควบคุมอัตราบิต ผลการทดสอบประสิทธิภาพของการเข้ารหัส จึงแบ่งออกเป็นหัวข้อต่าง ๆ ดังนี้

- ผลการทดสอบเมื่อปรับค่าถ่วงน้ำหนักบิตเป้าหมายในช่วง 0 ถึง 1 (0,0.1,0.2,...,1) เพื่อหาค่าถ่วงน้ำหนักบิตเป้าหมายที่เหมาะสม (Advanced Video Coding – Weight Target Bits Adjustment: AVC-W)
- ผลการทดสอบการกำหนดเงื่อนไขสำหรับการหาค่าควอนไทซ์พารามิเตอร์ (Advanced Video Coding – Conditional: AVC-C) โดยการหาจุดเริ่มเปลี่ยนบัพเฟออร์ที่เหมาะสมเมื่อกำหนดค่าถ่วงน้ำหนักแบบคงที่ (Advanced Video Coding – Conditional Weight Target Bits: AVC-CW) และแบบปรับตัวได้ (Advanced Video Coding – Conditional Adaptive Weight Target Bits: AVC-CAW)
- เปรียบเทียบการควบคุมอัตราที่นำเสนอกับการควบคุมอัตราการลงรหัสสัญญาณวีดิทัศน์ขั้นสูง

4.2.1 ผลการทดสอบเมื่อปรับค่า β ในช่วง 0 ถึง 1 (0, 0.1, 0.2, ..., 1)

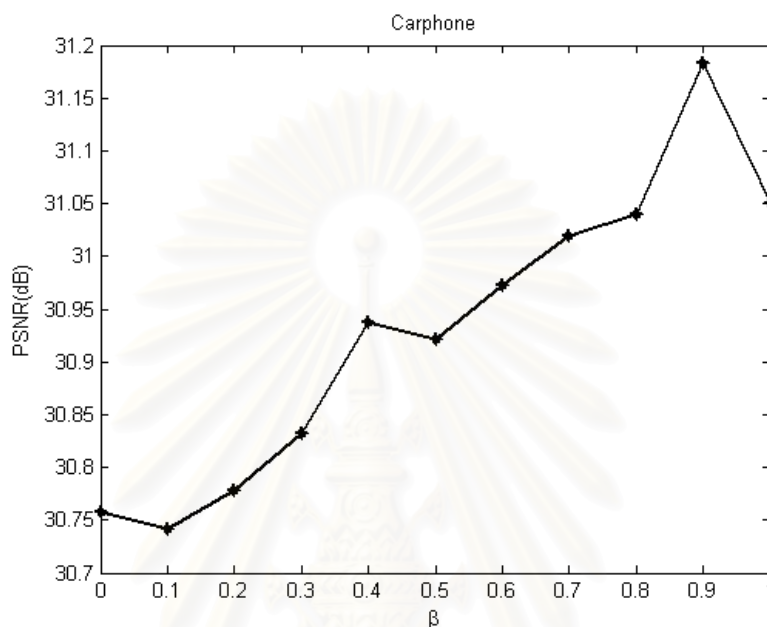
ในหัวข้อนี้เป็นการเปรียบเทียบผลการทดสอบการเข้ารหัส โดยปรับค่า β อยู่ในช่วง 0 ถึง 1 (0, 0.1, 0.2, ..., 1)

ผลการเปรียบเทียบการเข้ารหัสโดยปรับค่าถ่วงน้ำหนักบิตเป้าหมายของหลายๆลำดับภาพ แสดงดังรูปที่ 4.1 ถึง 4.5 และแสดงดังตารางที่ 4.2 ถึง 4.6 โดยสำหรับการควบคุมอัตราสำหรับการลงรหัสสัญญาณวีดิทัศน์ขั้นสูงคงที่กำหนดค่าถ่วงน้ำหนักบิตเป้าหมายคงที่เท่ากับ 0.5 โดยสังเกตได้ว่าการปรับค่าถ่วงน้ำหนักบิตเป้าหมาย มีผลต่อค่าคุณภาพสัญญาณวีดิทัศน์ และจำนวนเฟรมกระโดด โดยที่ค่าถ่วงน้ำหนักบิตเป้าหมายสูงขึ้นทำให้ค่าคุณภาพสัญญาณวีดิทัศน์เพิ่มขึ้น และจำนวนเฟรมกระโดดลดลง เนื่องจากในการควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูง ค่าถ่วงน้ำหนักบิตเป้าหมายถูกกำหนดให้มีค่าเท่ากับ 0.5 แต่เมื่อมีจำนวนบิตสะสมในบัพเฟออร์เพิ่มขึ้นจึงกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายเท่าเดิมไม่ได้ จึงต้องปรับเปลี่ยนค่าถ่วงน้ำหนักบิตเป้าหมายให้เพิ่มขึ้น ดังนั้นการกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายควรสามารถปรับค่าได้ เช่น ปรับค่าที่ระดับเฟรมตามค่ากิจกรรมการเคลื่อนที่ของเฟรมนั้นและจำนวนบิตสะสมในบัพเฟออร์ในขณะนั้น เป็นต้น

4.2.1.1 ผลการทดสอบกับลำดับภาพ Carphone

เมื่อทำการปรับค่าถ่วงน้ำหนักบิตเป้าหมายตั้งแต่ 0 ถึง 1 โดยเพิ่มขึ้นทีละ 0.1 แล้วพบว่า ค่าคุณภาพสัญญาณวีดิทัศน์เพิ่มขึ้น โดยค่าคุณภาพวีดิทัศน์สูงสุดที่ค่าถ่วงน้ำหนักบิต

เป้าหมายเท่ากับ 0.9 ดังแสดงดังรูปที่ 4.1 และจำนวนเฟรมกระโดดลดลงถึง 7 เฟรม เมื่อเทียบกับการกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายเท่ากับ 0.5 จำนวนเฟรมกระโดดและค่าคุณภาพสัญญาณวีดิทัศน์สามารถแสดงดังตารางที่ 4.2



รูปที่ 4.1 ค่าคุณภาพสัญญาณวีดิทัศน์เฉลี่ยที่ค่าถ่วงน้ำหนักบิตเป้าหมายต่าง ๆ ของลำดับภาพ Carphone

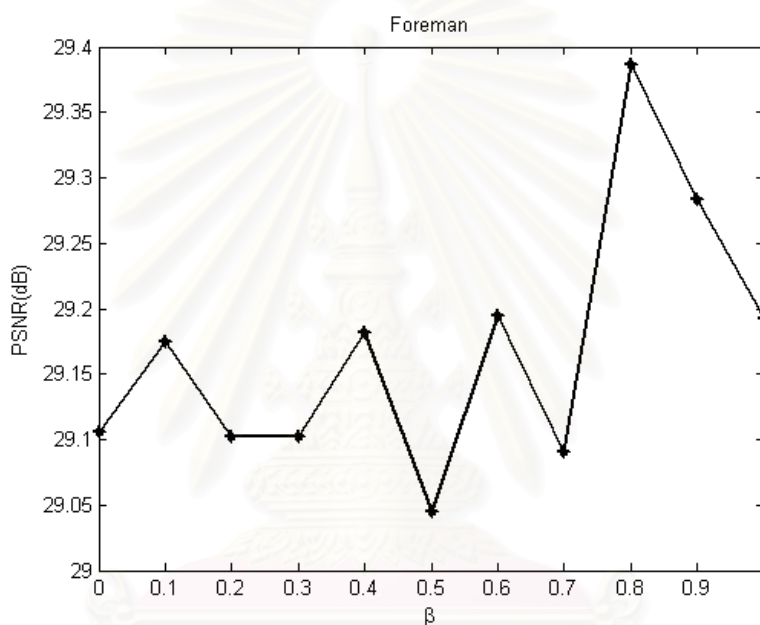
ตารางที่ 4.2 ค่าคุณภาพสัญญาณวีดิทัศน์และจำนวนเฟรมกระโดดที่ค่าถ่วงน้ำหนักบิตเป้าหมายต่าง ๆ ของลำดับภาพ Carphone

β	จำนวนเฟรมกระโดด	ค่าคุณภาพสัญญาณวีดิทัศน์ (dB)
0	20	30.75
0.1	20	30.74
0.2	19	30.77
0.3	17	30.83
0.4	15	30.93
0.5	15	30.92
0.6	12	30.97
0.7	12	31.01
0.8	12	31.03
0.9	8	31.18*
1	10	31.05

*คือ ค่าคุณภาพสัญญาณวีดิทัศน์สูงสุด

2.2.6 ผลการทดสอบกับลำดับภาพ Foreman

เมื่อทำการปรับค่าถ่วงน้ำหนักบิตเป้าหมายตั้งแต่ 0 ถึง 1 โดยเพิ่มขึ้นทีละ 0.1 แล้วพบว่าค่าคุณภาพสัญญาณวีดิทัศน์เพิ่มขึ้นโดยค่าคุณภาพวีดิทัศน์สูงสุดที่ค่าถ่วงน้ำหนักบิตเป้าหมายเท่ากับ 0.8 ดังแสดงดังรูปที่ 4.2 และจำนวนเฟรมกระโดดลดลงถึง 4 เฟรม เมื่อเทียบกับการกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายเท่ากับ 0.5 จำนวนเฟรมกระโดดและค่าคุณภาพสัญญาณวีดิทัศน์สามารถแสดงดังตารางที่ 4.3



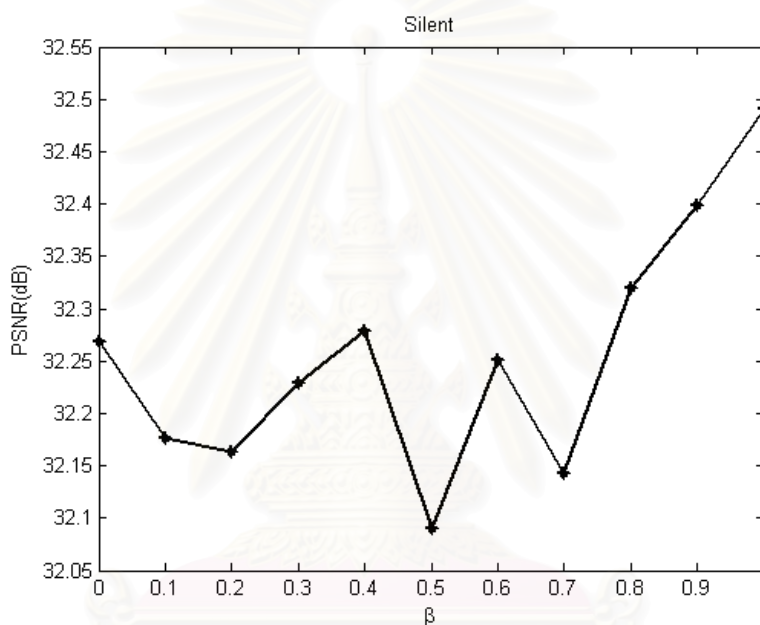
รูปที่ 4.2 ค่าคุณภาพสัญญาณวีดิทัศน์ที่ค่าถ่วงน้ำหนักบิตเป้าหมายต่างๆ
ของลำดับภาพ Foreman

ตารางที่ 4.3 ค่าคุณภาพสัญญาณวีดิทัศน์และจำนวนเฟรมกระโดดที่ค่าถ่วงน้ำหนักบิตเป้าหมาย
ต่าง ๆ ของลำดับภาพ Foreman

β	จำนวนเฟรมกระโดด	ค่าคุณภาพสัญญาณวีดิทัศน์ (dB)
0	22	29.10
0.1	23	29.17
0.2	21	29.10
0.3	21	29.10
0.4	20	29.18
0.5	21	29.04
0.6	20	29.19
0.7	21	29.09
0.8	17	29.38*
0.9	19	29.28
1	21	29.19

4.2.1.2 ผลการทดสอบกับลำดับภาพ Silent

เมื่อทำการปรับค่าถ่วงน้ำหนักบิตเป้าหมายตั้งแต่ 0 ถึง 1 โดยเพิ่มขึ้นทีละ 0.1 แล้วพบว่า ค่าคุณภาพสัญญาณวีดิทัศน์เพิ่มขึ้นโดยค่าคุณภาพวีดิทัศน์สูงสุดที่ค่าถ่วงน้ำหนักบิตเป้าหมายเท่ากับ 1.0 ดังแสดงดังรูปที่ 4.3 และจำนวนเฟรมกระโดดลดลงถึง 6 เฟรม เมื่อเทียบกับการกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายเท่ากับ 0.5 จำนวนเฟรมกระโดดและค่าคุณภาพสัญญาณวีดิทัศน์สามารถแสดงดังตารางที่ 4.4



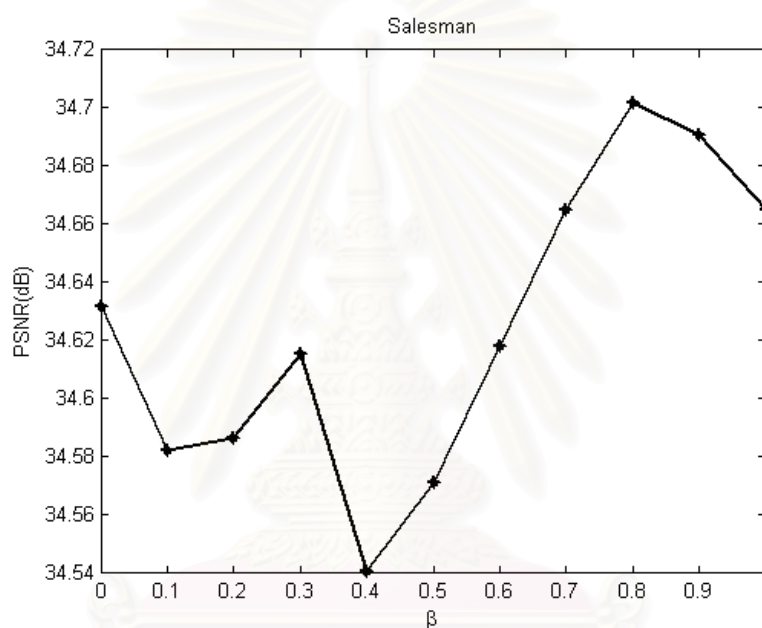
รูปที่ 4.3 ค่าคุณภาพสัญญาณวีดิทัศน์เฉลี่ยที่ค่าถ่วงน้ำหนักบิตเป้าหมายต่าง ๆ ของลำดับภาพ Silent

ตารางที่ 4.4 ค่าคุณภาพสัญญาณวีดิทัศน์และจำนวนเฟรมกระโดดที่ค่าถ่วงน้ำหนักบิตเป้าหมายต่าง ๆ ของลำดับภาพ Silent

β	จำนวนเฟรมกระโดด	ค่าคุณภาพสัญญาณวีดิทัศน์ (dB)
0	18	32.26
0.1	19	32.17
0.2	17	32.16
0.3	18	32.22
0.4	20	32.27
0.5	20	32.09
0.6	16	32.25
0.7	19	32.14
0.8	16	32.31
0.9	14	32.39
1	14	32.49*

4.2.1.4 ผลการทดสอบกับลำดับภาพ Salesman

เมื่อทำการปรับค่าถ่วงน้ำหนักบิตเป้าหมายตั้งแต่ 0 ถึง 1 โดยเพิ่มขึ้นทีละ 0.1 แล้วพบว่า ค่าคุณภาพสัญญาณวีดิทัศน์เพิ่มขึ้นโดยค่าคุณภาพวีดิทัศน์สูงสุดที่ค่าถ่วงน้ำหนักบิตเป้าหมาย เท่ากับ 1.0 ดังแสดงดังรูปที่ 4.4 และจำนวนเฟรมกระโดดลดลง 1 เฟรม เมื่อเทียบกับการกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายเท่ากับ 0.5 จำนวนเฟรมกระโดดและค่าคุณภาพสัญญาณวีดิทัศน์สามารถแสดงดังตารางที่ 4.5



รูปที่ 4.4 ค่าคุณภาพสัญญาณวีดิทัศน์เฉลี่ยที่ค่าถ่วงน้ำหนักบิตเป้าหมายต่าง ๆ ของลำดับภาพ Salesman

ตารางที่ 4.5 ค่าคุณภาพสัญญาณวีดิทัศน์และจำนวนเฟรมกระโดดที่ค่าถ่วงน้ำหนักบิตเป้าหมายต่าง ๆ ของลำดับภาพ Salesman

β	จำนวนเฟรมกระโดด	ค่าคุณภาพสัญญาณวีดิทัศน์ (dB)
0	16	34.63
0.1	18	34.58
0.2	14	34.58
0.3	14	34.61
0.4	19	34.54
0.5	16	34.57
0.6	15	34.61
0.7	15	34.66
0.8	15	34.70*
0.9	14	34.69
1	13	34.66

4.2.2 ผลการทดสอบการกำหนดเงื่อนไขสำหรับการหาค่าควอนไทล์พารามิเตอร์โดยการหาค่าจุดเริ่มเปลี่ยนบัพเฟอร์ที่เหมาะสมเมื่อกำหนดค่าถ่วงน้ำหนักแบบคงที่และแบบปรับตัวได้

ในหัวข้อนี้ทำการทดสอบเพื่อหาค่าจุดเริ่มเปลี่ยนบัพเฟอร์ที่เหมาะสม โดยปรับค่าจุดเริ่มเปลี่ยนบัพเฟอร์ เท่ากับ 100 125 150 และ 175 หน่วย (ในกรณีนี้จำนวนบิตเป้าหมายในระดับเฟรมมีค่าน้อยกว่าศูนย์ จำนวนบิตสะสมในบัพเฟอร์ในขณะนี้ คือ ผลรวมของจำนวนบิตที่เข้ารหัสตั้งแต่หน่วยพื้นฐานย่อยที่ 1 ถึง หน่วยพื้นฐานย่อยปัจจุบัน กับ จำนวนบิตสะสมในบัพเฟอร์ของเฟรมก่อนหน้า) โดยการเลือกค่าจุดเริ่มเปลี่ยนบัพเฟอร์ที่เหมาะสมค่านึงถึงจำนวนบิตสะสมในบัพเฟอร์กล่าว คือ ถ้าเลือกค่าจุดเริ่มเปลี่ยนบัพเฟอร์น้อยเกินไปอาจเกิดกรณีไม่มีข้อมูลในบัพเฟอร์ ซึ่งเป็นการใช้ช่องสัญญาณอย่างไม่มีประสิทธิภาพ และในทางกลับกันถ้าเลือกค่าจุดเริ่มเปลี่ยนบัพเฟอร์มากเกินไป จะทำให้จำนวนบิตสะสมในบัพเฟอร์เกินกว่าค่าสูงสุดของบัพเฟอร์ ทำให้เกิดการกระโดดข้ามเฟรม ทำการทดสอบเปรียบเทียบ โดยกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายคงที่ และการกำหนดค่าถ่วงน้ำหนักแบบปรับตัว ที่ค่าจุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ

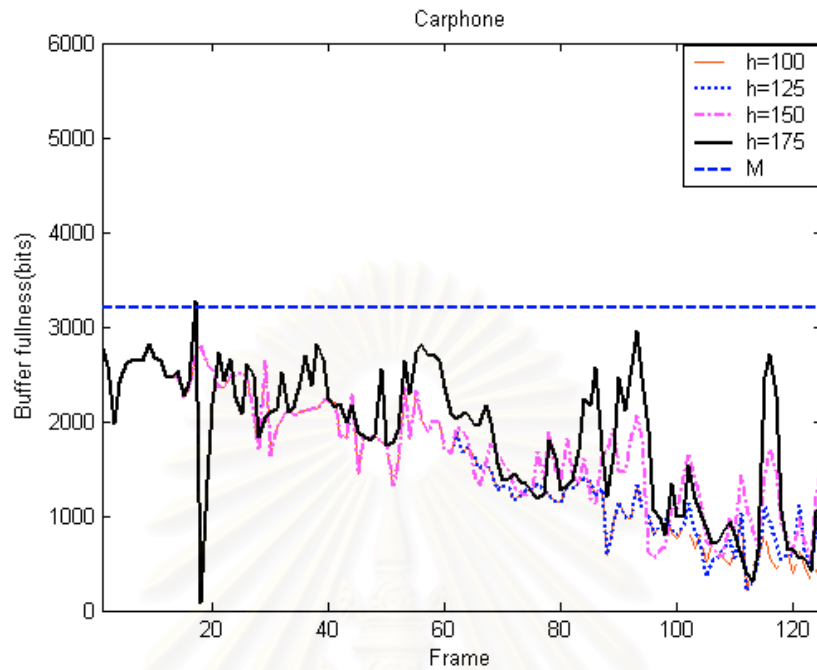
4.2.2.1 ผลการทดสอบกับลำดับภาพ Carphone

ก) ค่าถ่วงน้ำหนักบิตเป้าหมายคงที่

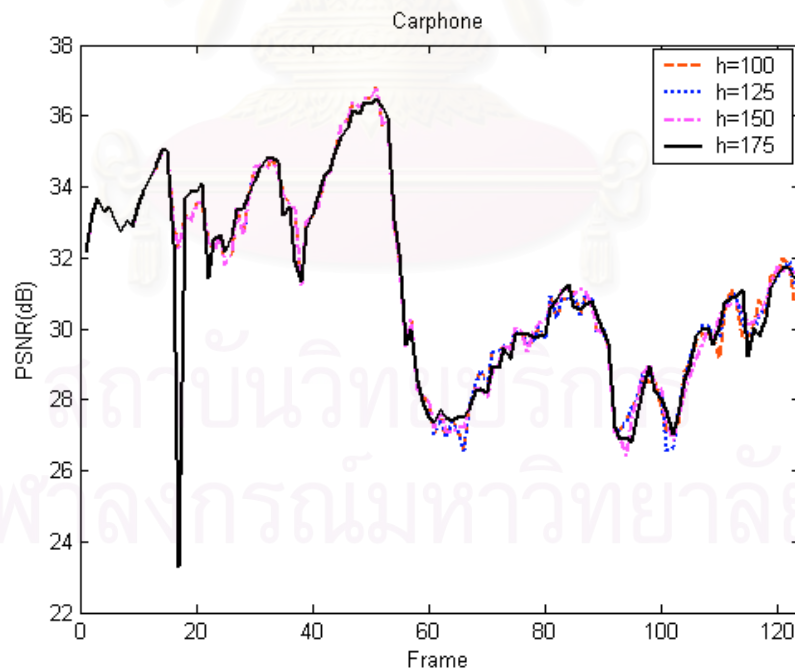
โดยกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายคงที่เท่ากับ 0.9 จากรูปที่ 4.5 และ 4.6 คือ จำนวนบิตสะสมในบัพเฟอร์และค่าคุณภาพสัญญาณวีดิทัศน์แต่ละเฟรม ที่ค่าจุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ ตามลำดับ

ข) ค่าถ่วงน้ำหนักบิตเป้าหมายปรับตัวได้

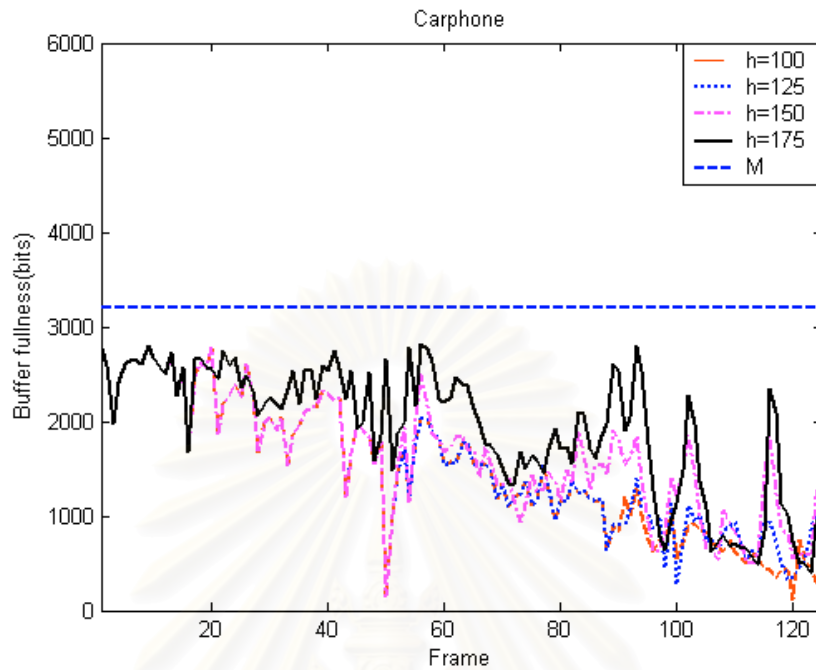
ค่า $\beta_{initial}$ กำหนดเท่ากับ 0.9 ค่า β ปรับเพิ่มและลดลงอีก 0.1 โดยค่า Threshold เท่ากับ 200 จากรูปที่ 4.7 และ 4.8 คือ จำนวนบิตสะสมในบัพเฟอร์และค่าคุณภาพสัญญาณวีดิทัศน์แต่ละเฟรมที่ค่าจุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ ตามลำดับ รูปที่ 4.9 แสดงการปรับตัวของค่าถ่วงน้ำหนักบิตเป้าหมายเมื่อค่าจุดเริ่มเปลี่ยนบัพเฟอร์ เท่ากับ 175 หน่วย



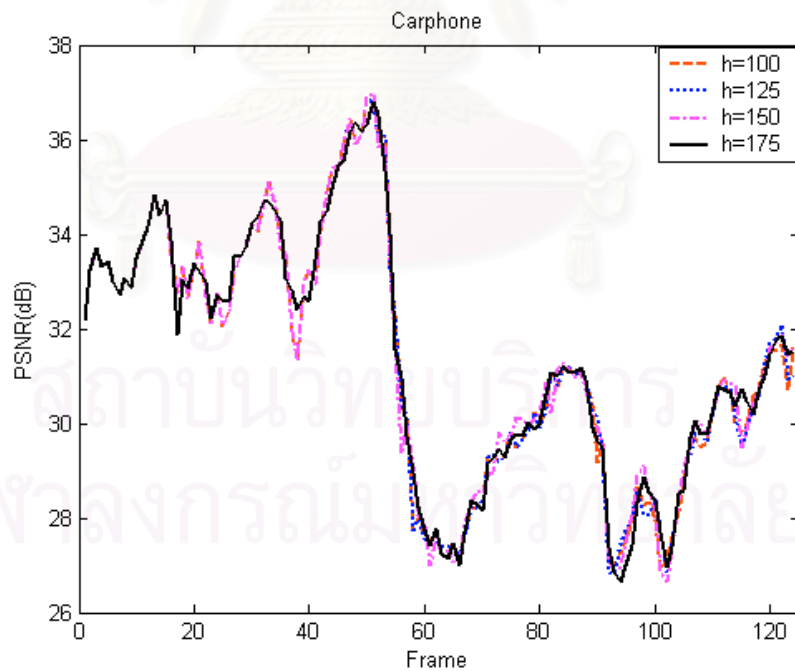
รูปที่ 4.5 จำนวนบิตสะสมในบัฟเฟอร์ของลำดับภาพ Carphone ที่จุดเริ่มเปลี่ยนบัฟเฟอร์ต่าง ๆ เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายคงที่



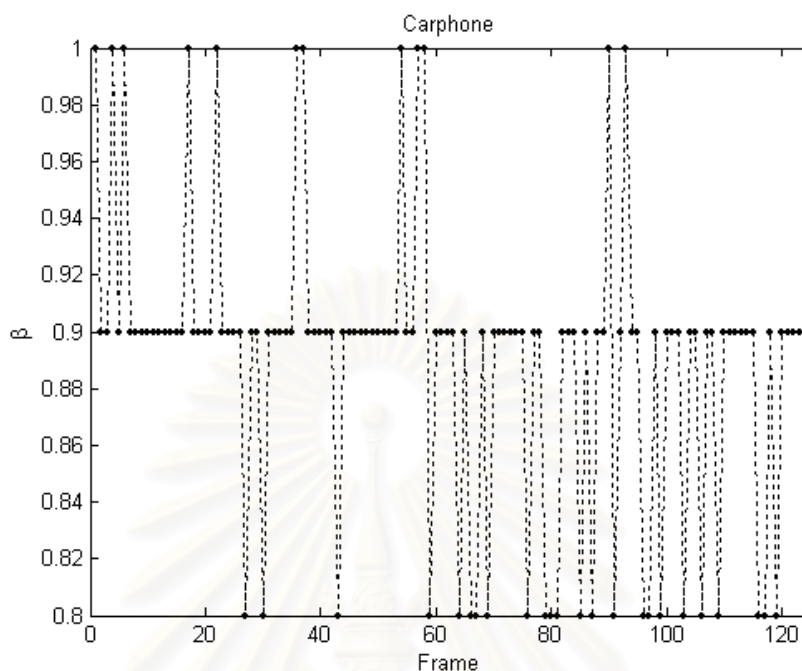
รูปที่ 4.6 ค่าคุณภาพสัญญาณวีดิทัศน์ของลำดับภาพ Carphone ที่จุดเริ่มเปลี่ยนบัฟเฟอร์ต่างๆ เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายคงที่



รูปที่ 4.7 จำนวนบิตสะสมในบัฟเฟอร์ของลำดับภาพ Carphone ที่จุดเริ่มเปลี่ยนบัฟเฟอร์ต่าง ๆ เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายปรับตัวได้



รูปที่ 4.8 ค่าคุณภาพสัญญาณวิดีโอที่ศรัณย์ของลำดับภาพ Carphone ที่จุดเริ่มเปลี่ยนบัฟเฟอร์ต่าง ๆ เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายปรับตัวได้



รูปที่ 4.9 ค่าถ่วงน้ำหนักบิตเป้าหมายของลำดับภาพ Carphone

จากรูปที่ 4.5 - 4.8 สามารถสรุปได้ดังตารางที่ 4.6 เมื่อเปรียบเทียบที่จุดเริ่มเปลี่ยนบัพเฟอร์ต่างๆพบว่า เมื่อปรับค่าถ่วงน้ำหนักบิตเป้าหมายคงที่ ค่าจุดเริ่มเปลี่ยนบัพเฟอร์ที่ 150 หน่วย ให้คุณภาพสัญญาณวีดิทัศน์ดีที่สุด สำหรับการกำหนดค่าบิตเป้าหมายแบบปรับตัวได้ ค่าจุดเริ่มเปลี่ยนบัพเฟอร์ที่ 175 หน่วย ให้ค่าคุณภาพสัญญาณวีดิทัศน์ดีที่สุด

เมื่อเปรียบเทียบระหว่างการกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายคงที่กับแบบปรับตัวได้ที่จุดเริ่มเปลี่ยนบัพเฟอร์เดียวกัน พบว่า ที่จุดเริ่มเปลี่ยนบัพเฟอร์ที่ 175 หน่วย การกำหนดค่าถ่วงน้ำหนักแบบปรับตัวสามารถลดจำนวนเฟรมได้ 100 % และมีค่าคุณภาพสัญญาณวีดิทัศน์ดีกว่า

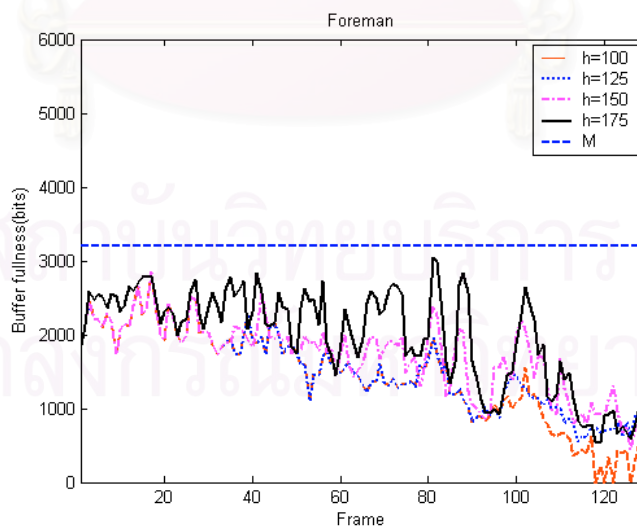
4.2.2.2 ผลการทดสอบกับลำดับภาพ Foreman

ก) ค่าถ่วงน้ำหนักบิตเป้าหมายคงที่

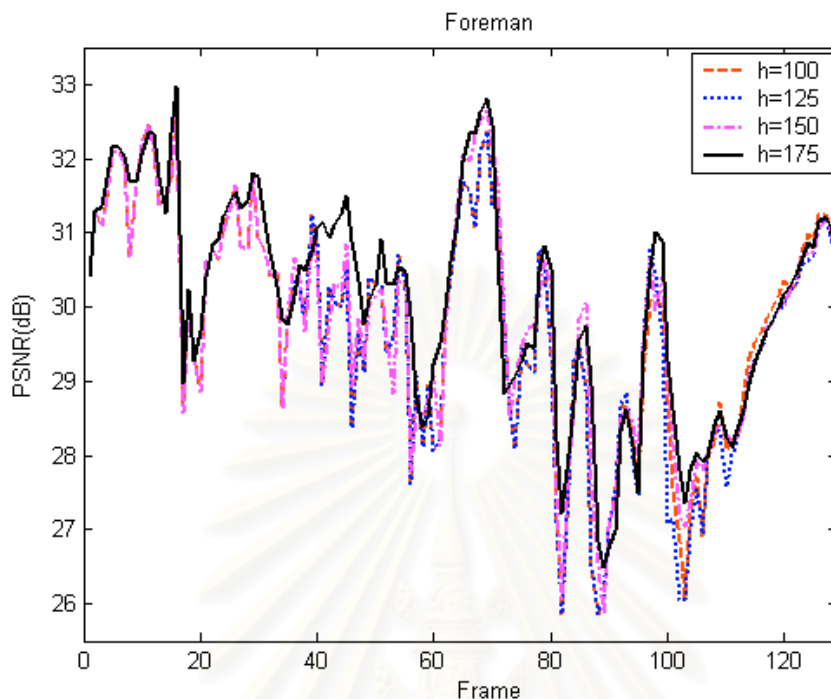
โดยกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายคงที่เท่ากับ 0.9 จากรูปที่ 4.10 และ 4.11 คือ จำนวนบิตสะสมในบัพเฟอร์และค่าคุณภาพสัญญาณวีดิทัศน์แต่ละเฟรมที่ค่าจุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ ตามลำดับ

ตารางที่ 4.6 จำนวนเฟรมกระโดด จำนวนบิตสะสมในบัฟเฟอร์ และ ค่าคุณภาพสัญญาณวีดิทัศน์ ที่ค่าจุดเริ่มเปลี่ยนบัฟเฟอร์ต่าง ๆ ของลำดับภาพ Carphone เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายแบบคงที่และแบบปรับตัวได้

h	ค่าถ่วงน้ำหนักบิตเป้าหมายแบบคงที่				ค่าถ่วงน้ำหนักบิตเป้าหมายแบบปรับตัวได้			
	จำนวนเฟรมกระโดด	ความเต็มของบัฟเฟอร์		ค่าคุณภาพสัญญาณวีดิทัศน์ (dB)	จำนวนเฟรมกระโดด	ความเต็มของบัฟเฟอร์		ค่าคุณภาพสัญญาณวีดิทัศน์ (dB)
		น้อยเกินเก็บ (under flow)	มากเกินเก็บ (over flow)			น้อยเกินเก็บ (under flow)	มากเกินเก็บ (over flow)	
100	0	ไม่	ไม่	31.44	0	ไม่	ไม่	31.44
125	0	ไม่	ไม่	31.45	0	ไม่	ไม่	31.45
150	0	ไม่	ไม่	31.44	0	ไม่	ไม่	31.47
175	1	ไม่	ใช่	31.28	0	ไม่	ไม่	31.49



รูปที่ 4.10 จำนวนบิตสะสมในบัฟเฟอร์ของลำดับภาพ Foreman ที่จุดเริ่มเปลี่ยนบัฟเฟอร์ต่าง ๆ เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายคงที่



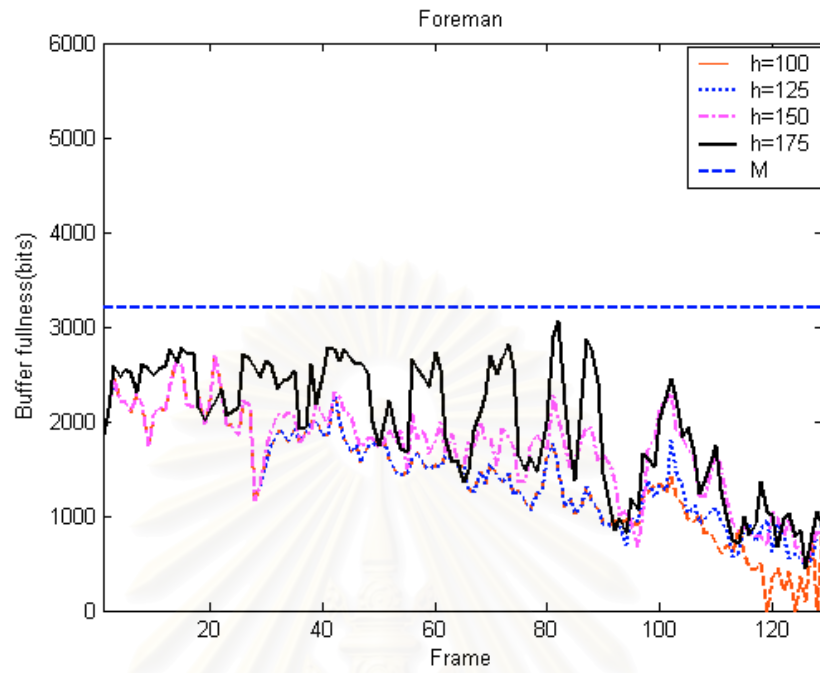
รูปที่ 4.11 ค่าคุณภาพสัญญาณวิดีโอที่ค้นหาลำดับภาพ Foreman ที่จุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายคงที่

(ข) ค่าถ่วงน้ำหนักบิตเป้าหมายปรับตัวได้

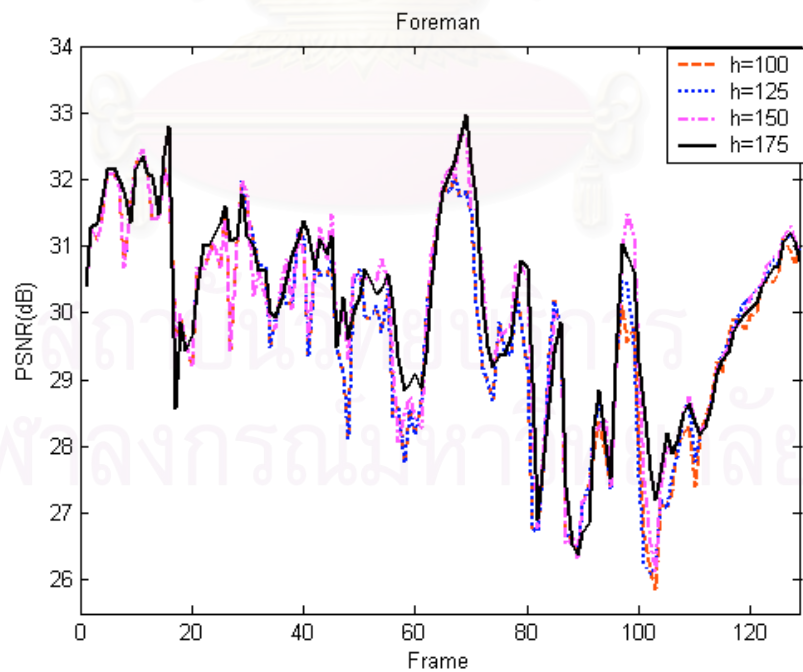
ค่า $\beta_{initial}$ กำหนดเท่ากับ 0.8 ค่า β ปรับเพิ่มและลดลงอีก 0.1 โดยค่า Threshold เท่ากับ 400 จากรูปที่ 4.12 และ 4.13 คือ จำนวนบิตสะสมในบัพเฟอร์ และค่าคุณภาพสัญญาณวิดีโอที่ค้นหแต่ละเฟรม ที่ค่าจุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ ตามลำดับ รูปที่ 4.14 แสดงการปรับตัวของค่าถ่วงน้ำหนักบิตเป้าหมายเมื่อค่าจุดเริ่มเปลี่ยนบัพเฟอร์ เท่ากับ 175 หน่วย

จากรูปที่ 4.10 - 4.13 สามารถสรุปได้ดังตารางที่ 4.7 เมื่อเปรียบเทียบที่จุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ พบว่า ค่าจุดเริ่มเปลี่ยนบัพเฟอร์ที่ 175 หน่วย ให้คุณภาพสัญญาณวิดีโอที่ค้นหดีที่สุด ทั้งการกำหนดค่าบิตเป้าหมายแบบคงที่และแบบปรับตัวได้

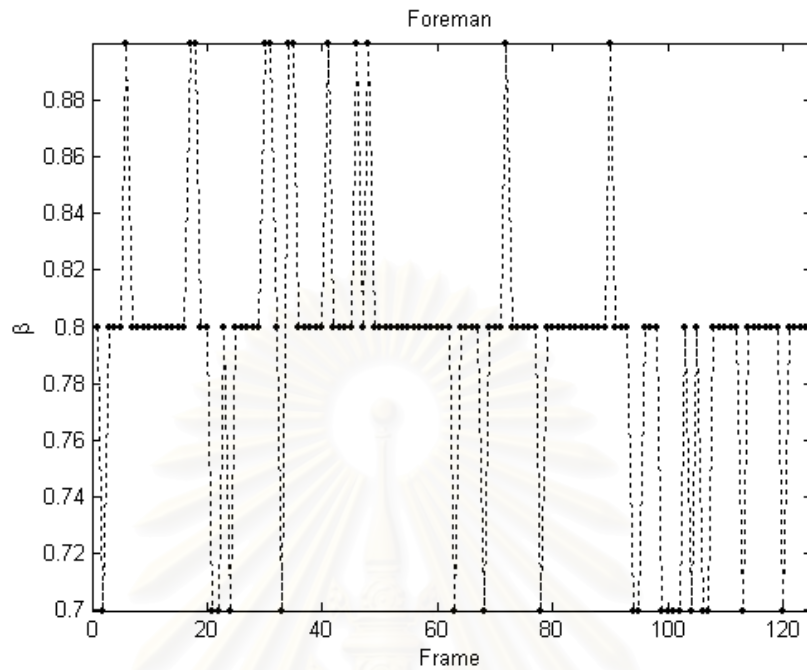
เมื่อเปรียบเทียบระหว่างการกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายคงที่กับแบบปรับตัวได้ที่จุดเริ่มเปลี่ยนบัพเฟอร์เดียวกัน พบว่า ที่จุดเริ่มเปลี่ยนบัพเฟอร์ที่ 175 หน่วย การกำหนดค่าถ่วงน้ำหนักแบบปรับตัว มีค่าคุณภาพสัญญาณวิดีโอที่ค้นหที่ดีกว่า



รูปที่ 4.12 จำนวนบิตสะสมในบัฟเฟอร์ของลำดับภาพ Foreman ที่จุดเริ่มเปลี่ยนบัฟเฟอร์ต่าง ๆ เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายปรับตัวได้



รูปที่ 4.13 ค่าคุณภาพสัญญาณวิดีโอที่ต้นขงลำดับภาพ Foreman ที่จุดเริ่มเปลี่ยนบัฟเฟอร์ต่าง ๆ เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายปรับตัวได้



รูปที่ 4.14 ค่าถ่วงน้ำหนักบิตเป้าหมายของลำดับภาพ Foreman

ตารางที่ 4.7 จำนวนเฟรมกระโดด จำนวนบิตสะสมในบัฟเฟอร์ และ ค่าคุณภาพสัญญาณวีดิทัศน์
ที่ค่าจุดเริ่มเปลี่ยนบัฟเฟอร์ ต่างๆ ของลำดับภาพ Foreman เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายแบบคงที่และแบบปรับตัวได้

h	ค่าถ่วงน้ำหนักบิตเป้าหมายแบบคงที่				ค่าถ่วงน้ำหนักบิตเป้าหมายแบบปรับตัวได้			
	จำนวนเฟรมกระโดด	ความเต็มของบัฟเฟอร์		ค่าคุณภาพสัญญาณวีดิทัศน์ (dB)	จำนวนเฟรมกระโดด	ความเต็มของบัฟเฟอร์		ค่าคุณภาพสัญญาณวีดิทัศน์ (dB)
		น้อยเกินเก็บ (under flow)	มากเกินเก็บ (over flow)			น้อยเกินเก็บ (under flow)	มากเกินเก็บ (over flow)	
100	0	ใช่	ไม่	29.93	0	ใช่	ไม่	29.93
125	0	ไม่	ไม่	29.89	0	ไม่	ไม่	29.96
150	0	ไม่	ไม่	30.03	0	ไม่	ไม่	30.16
175	0	ไม่	ไม่	30.22	0	ไม่	ไม่	30.23

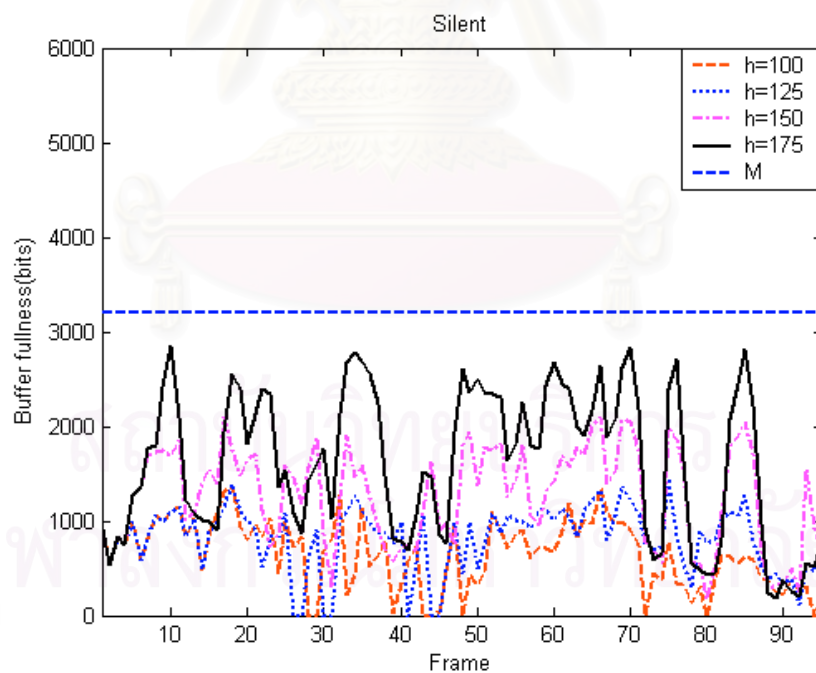
4.2.2.2 ผลการทดสอบกับลำดับภาพ Silent

ก) ค่าถ่วงน้ำหนักบิตเป้าหมายคงที่

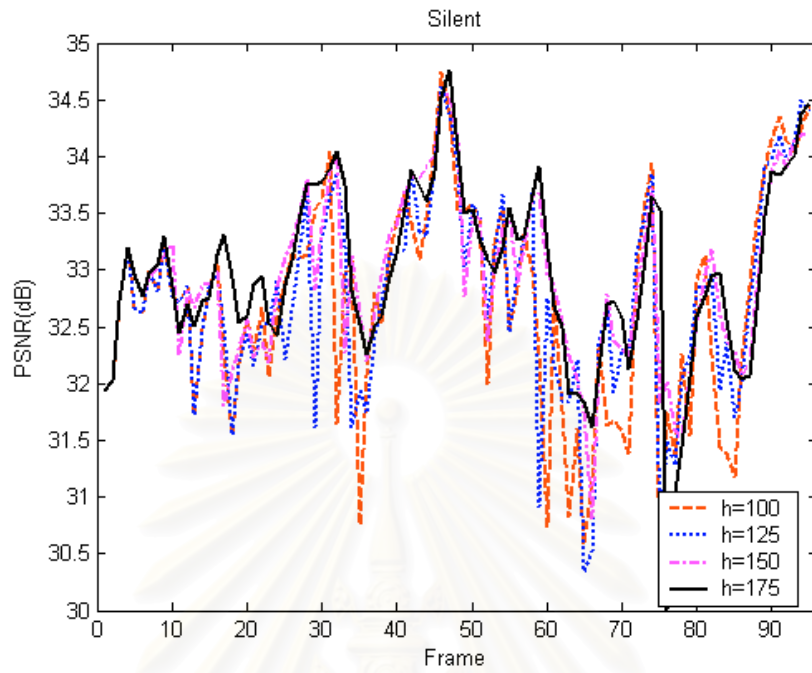
โดยกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายคงที่เท่ากับ 1.0 จากรูปที่ 4.15 และ 4.16 คือ จำนวนบิตสะสมในบัฟเฟอร์ และค่าคุณภาพสัญญาณวีดิทัศน์แต่ละเฟรมที่ค่าจุดเริ่มเปลี่ยนบัฟเฟอร์ต่าง ๆ ตามลำดับ

ข) ค่าถ่วงน้ำหนักบิตเป้าหมายปรับตัวได้

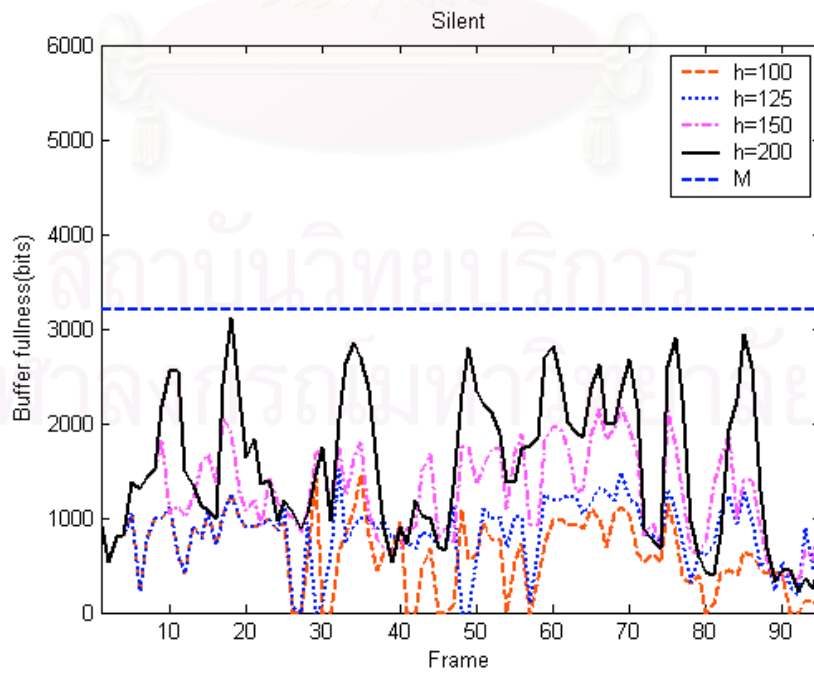
สำหรับลำดับภาพ Silent ค่า $\beta_{initial}$ กำหนดเท่ากับ 1.0 ค่า β ปรับลดลงอีก 0.1 โดยค่า Threshold เท่ากับ 300 รูปที่ 4.17 และ 4.18 คือ จำนวนบิตสะสมในบัฟเฟอร์ และค่าคุณภาพสัญญาณวีดิทัศน์แต่ละเฟรมที่ค่าจุดเริ่มเปลี่ยนบัฟเฟอร์ต่าง ๆ ตามลำดับ รูปที่ 4.19 แสดงการปรับตัวของค่าถ่วงน้ำหนักบิตเป้าหมายเมื่อค่าจุดเริ่มเปลี่ยนบัฟเฟอร์เท่ากับ 175 หน่วย



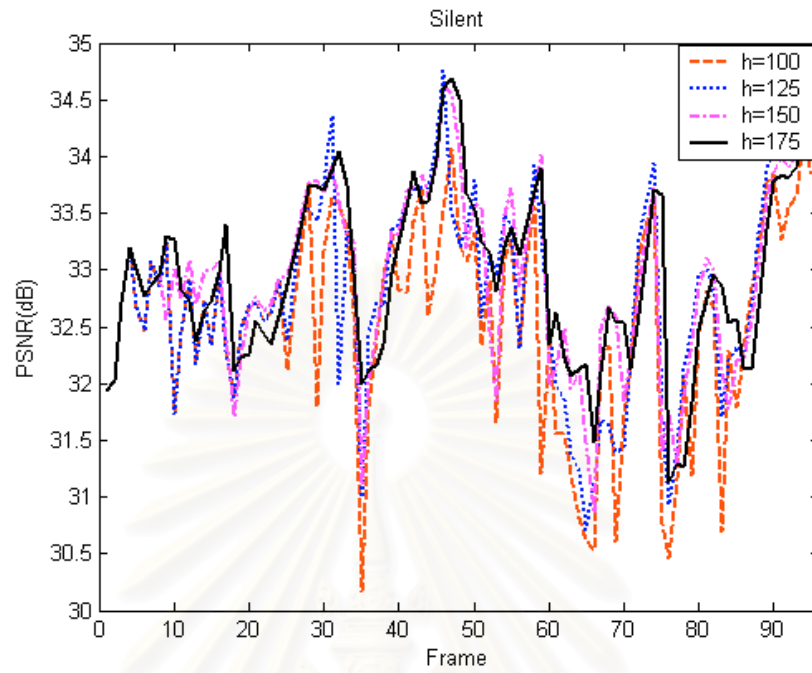
รูปที่ 4.15 จำนวนบิตสะสมในบัฟเฟอร์ของลำดับภาพ Silent ที่จุดเริ่มเปลี่ยนบัฟเฟอร์ต่าง ๆ เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายคงที่



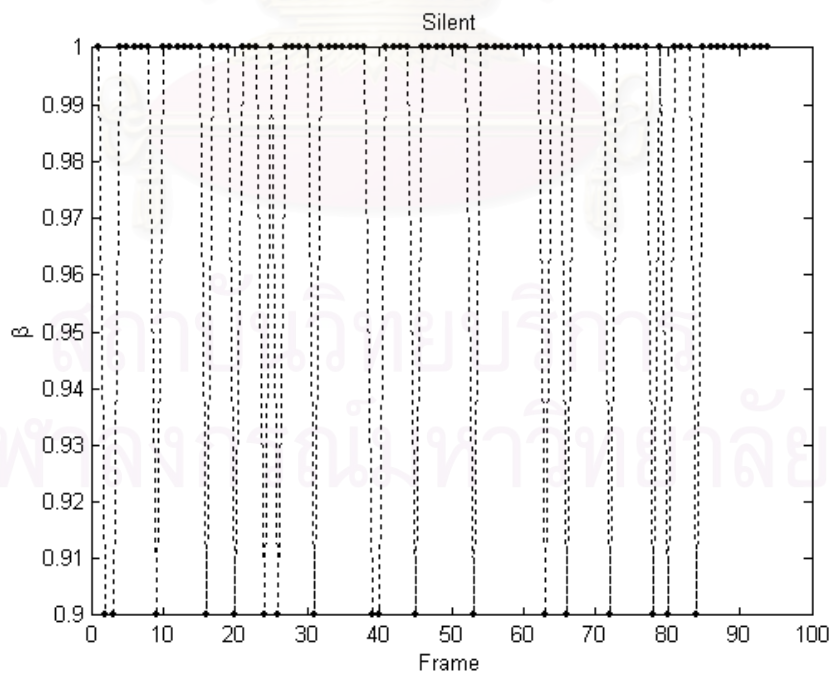
รูปที่ 4.16 ค่าคุณภาพสัญญาณวีดิทัศน์ของลำดับภาพ Silent ที่จุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายปรับตัวคงที่



รูปที่ 4.17 จำนวนบิตสะสมในบัพเฟอร์ของลำดับภาพ Silent ที่จุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายปรับตัวได้



รูปที่ 4.18 ค่าคุณภาพสัญญาณวีดิทัศน์ของลำดับภาพ Silent ที่จุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายปรับตัวได้



รูปที่ 4.19 ค่าถ่วงน้ำหนักบิตเป้าหมายของลำดับภาพ Silent

จากรูปที่ 4.15 ถึง 4.18 สามารถสรุปได้ดังตารางที่ 4.8 เมื่อเปรียบเทียบที่จุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ พบว่า ค่าจุดเริ่มเปลี่ยนบัพเฟอร์ที่ 175 หน่วย ให้ค่าคุณภาพสัญญาณวีดิทัศน์ที่ดีที่สุด ทั้งการกำหนดค่าบิตเป้าหมายแบบคงที่และแบบปรับตัวได้

เมื่อเปรียบเทียบระหว่างการกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายคงที่กับแบบปรับตัวได้ที่จุดเริ่มเปลี่ยนบัพเฟอร์เดียวกัน พบว่าที่จุดเริ่มเปลี่ยนบัพเฟอร์ที่ 175 หน่วย การกำหนดค่าถ่วงน้ำหนักแบบปรับตัว มีค่าคุณภาพสัญญาณวีดิทัศน์ที่ดีกว่า

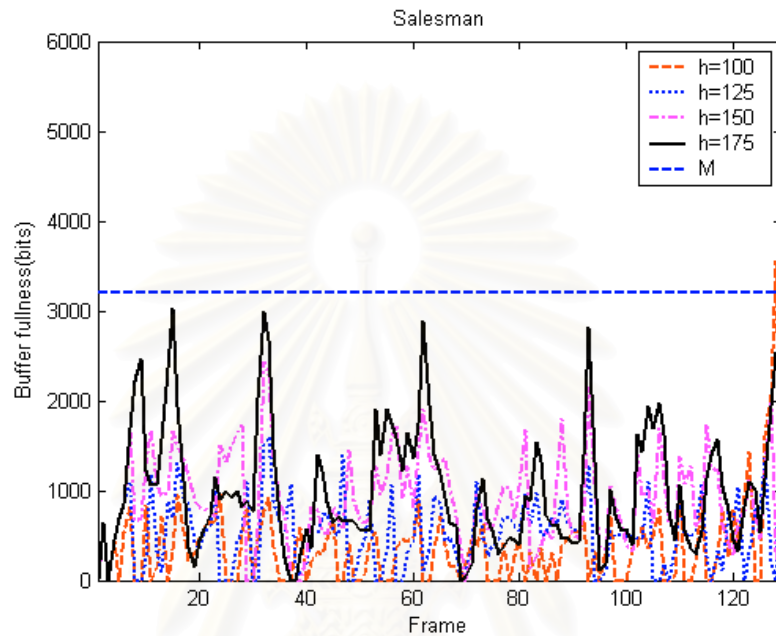
ตารางที่ 4.8 จำนวนเฟรมกระโดด จำนวนบิตสะสมในบัพเฟอร์ และค่าคุณภาพสัญญาณวีดิทัศน์ที่ค่าจุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ ของลำดับภาพ Silent เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายแบบคงที่และแบบปรับตัวได้

h	ค่าถ่วงน้ำหนักบิตเป้าหมายแบบคงที่				ค่าถ่วงน้ำหนักบิตเป้าหมายแบบปรับตัวได้			
	จำนวนเฟรมกระโดด	ความเต็มของบัพเฟอร์		ค่าคุณภาพสัญญาณวีดิทัศน์ (dB)	จำนวนเฟรมกระโดด	ความเต็มของบัพเฟอร์		ค่าคุณภาพสัญญาณวีดิทัศน์ (dB)
		น้อยเกินเก็บ (under flow)	มากเกินเก็บ (over flow)			น้อยเกินเก็บ (under flow)	มากเกินเก็บ (over flow)	
100	0	ใช่	ไม่	32.73	0	ใช่	ไม่	32.59
125	0	ใช่	ไม่	32.79	0	ใช่	ไม่	32.84
150	0	ไม่	ไม่	32.95	0	ไม่	ไม่	32.94
175	0	ไม่	ไม่	33.00	0	ไม่	ไม่	33.01

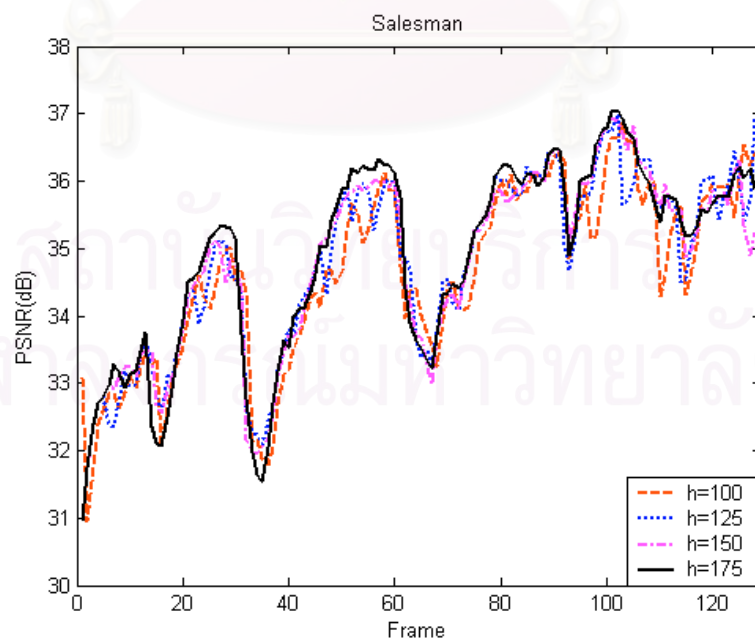
4.2.2.4 ผลการทดสอบกับลำดับภาพ Salesman

ก) ค่าถ่วงน้ำหนักบิตเป้าหมายคงที่

โดยกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายคงที่เท่ากับ 0.9 จากรูปที่ 4.20 และ 4.21 คือ จำนวนบิตสะสมในบัฟเฟอร์ และค่าคุณภาพสัญญาณวีดิทัศน์แต่ละเฟรมตามลำดับ ที่ค่าจุดเริ่มเปลี่ยนบัฟเฟอร์ต่าง ๆ



รูปที่ 4.20 จำนวนบิตสะสมในบัฟเฟอร์ของลำดับภาพ Salesman ที่จุดเริ่มเปลี่ยนบัฟเฟอร์ต่าง ๆ เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายคงที่



รูปที่ 4.21 ค่าคุณภาพสัญญาณวีดิทัศน์ของลำดับภาพ Salesman ที่จุดเริ่มเปลี่ยนบัฟเฟอร์ต่าง ๆ เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายคงที่

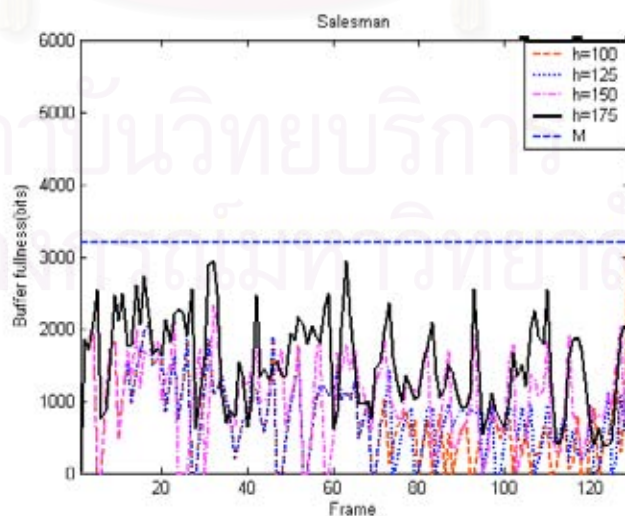
ข) ค่าถ่วงน้ำหนักบิตเป้าหมายปรับตัวได้

สำหรับลำดับภาพ Salesman ค่า $\beta_{initial}$ กำหนดเท่ากับ 0.9 ค่า β ปรับลดลงอีก 0.1 โดยค่า Threshold เท่ากับ 70 รูปที่ 4.22 และ 4.23 คือ จำนวนบิตสะสมในบัฟเฟอร์และค่าคุณภาพสัญญาณวีดิทัศน์แต่ละเฟรมที่ค่าจุดเริ่มเปลี่ยนบัฟเฟอร์ต่าง ๆ ตามลำดับ รูปที่ 4.24 แสดงการปรับตัวของค่าถ่วงน้ำหนักบิตเป้าหมายเมื่อค่าจุดเริ่มเปลี่ยนบัฟเฟอร์เท่ากับ 175 หน่วย

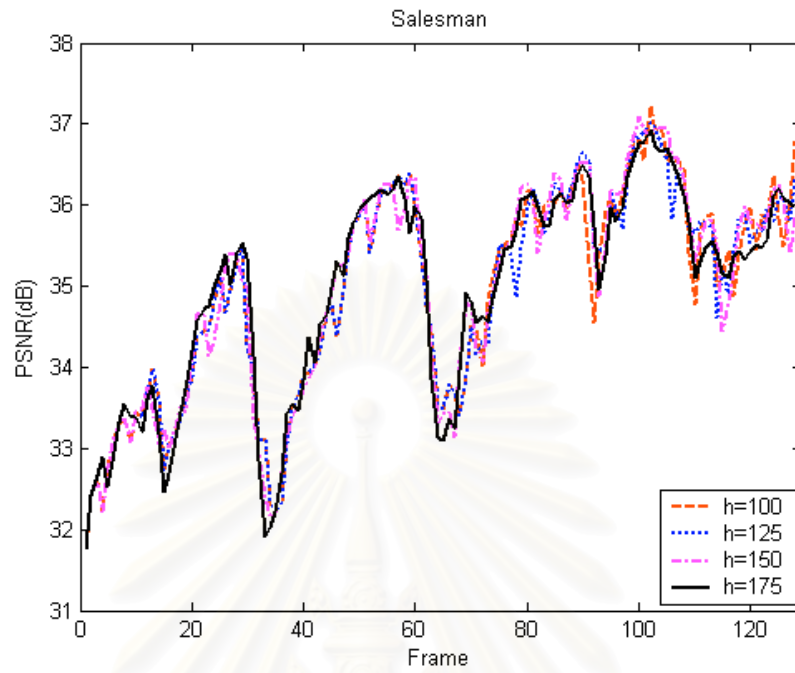
จากรูปที่ 4.20 ถึง 4.23 สามารถสรุปดังตารางที่ 4.9 เมื่อเปรียบเทียบที่จุดเริ่มเปลี่ยนบัฟเฟอร์ต่างๆ พบว่า ค่าจุดเริ่มเปลี่ยนบัฟเฟอร์ที่ 175 หน่วย ให้ค่าคุณภาพสัญญาณวีดิทัศน์ที่ดีที่สุด ทั้งการกำหนดค่าบิตเป้าหมายแบบคงที่และแบบปรับตัวได้

เมื่อเปรียบเทียบระหว่างการกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายคงที่กับแบบปรับตัวได้ที่จุดเริ่มเปลี่ยนบัฟเฟอร์เดียวกันพบว่าที่จุดเริ่มเปลี่ยนบัฟเฟอร์ที่ 175 หน่วย การกำหนดค่าถ่วงน้ำหนักแบบปรับตัว ค่าคุณภาพสัญญาณวีดิทัศน์ดีกว่า อีกทั้งการกำหนดบิตเป้าหมายคงที่ เกิดกรณีที่จำนวนบิตสะสมน้อยเกินไปทั้ง 4 กรณี ซึ่งเป็นการใช้ช่องสัญญาณอย่างไม่มีประสิทธิภาพ

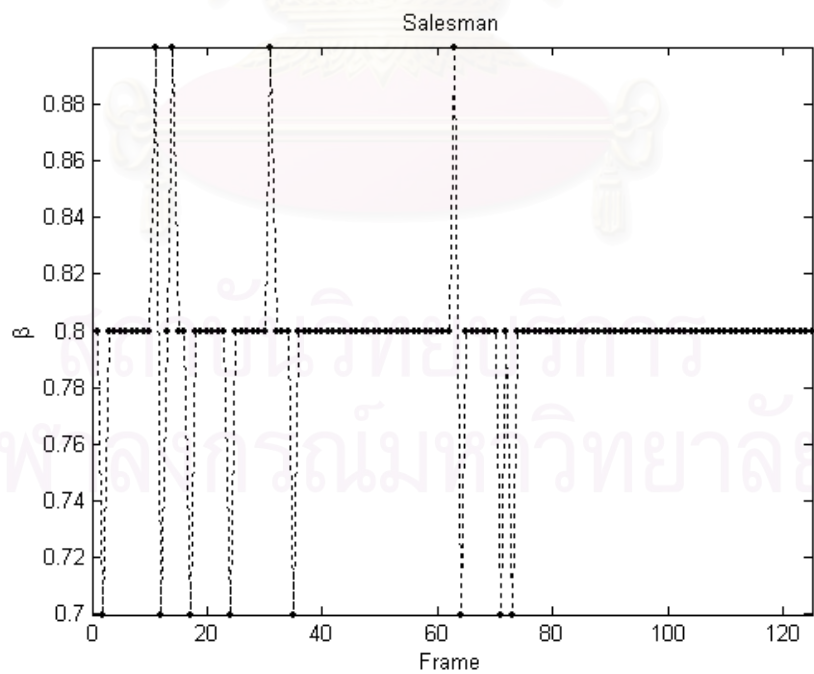
จากตารางที่ 4.5-4.8 สามารถสรุปได้ว่าการกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายสามารถให้ประสิทธิภาพดีกว่าการกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายแบบคงที่และค่าจุดเริ่มเปลี่ยนที่เหมาะสมคือ 175 หน่วย



รูปที่ 4.22 จำนวนบิตสะสมในบัฟเฟอร์ของลำดับภาพ Salesman ที่จุดเริ่มเปลี่ยนบัฟเฟอร์ต่าง ๆ เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายปรับตัวได้



รูปที่ 4.23 ค่าคุณภาพสัญญาณวิดีโอที่คืนของลำดับภาพ Salesman ที่จุดเริ่มเปลี่ยนบัพเฟออร์ต่างๆ เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายปรับตัวได้



รูปที่ 4.24 ค่าถ่วงน้ำหนักบิตเป้าหมายของลำดับภาพ Salesman

ตารางที่ 4.9 จำนวนเฟรมกระโดด จำนวนบิตสะสมในบัพเฟอร์ และ ค่าคุณภาพสัญญาณวีดิทัศน์ ที่ค่าจุดเริ่มเปลี่ยนบัพเฟอร์ต่าง ๆ ของลำดับภาพ Salesman เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้า หมายแบบคงที่และแบบปรับตัวได้

h	ค่าถ่วงน้ำหนักบิตเป้าหมายแบบคงที่				ค่าถ่วงน้ำหนักบิตเป้าหมายแบบปรับตัวได้			
	จำนวนเฟรมกระโดด	ความเต็มของบัพเฟอร์		ค่าคุณภาพสัญญาณวีดิทัศน์ (dB)	จำนวนเฟรมกระโดด	ความเต็มของบัพเฟอร์		ค่าคุณภาพสัญญาณวีดิทัศน์ (dB)
		น้อยเกินเก็บ (under flow)	มากเกินไปเก็บ (over flow)			น้อยเกินเก็บ (under flow)	มากเกินไปเก็บ (over flow)	
100	1	ใช่	ใช่	34.50	0	ใช่	ไม่	34.85
125	0	ใช่	ไม่	34.7	0	ใช่	ไม่	34.84
150	0	ใช่	ไม่	34.73	0	ใช่	ไม่	34.86
175	0	ใช่	ไม่	34.79	0	ไม่	ไม่	34.87

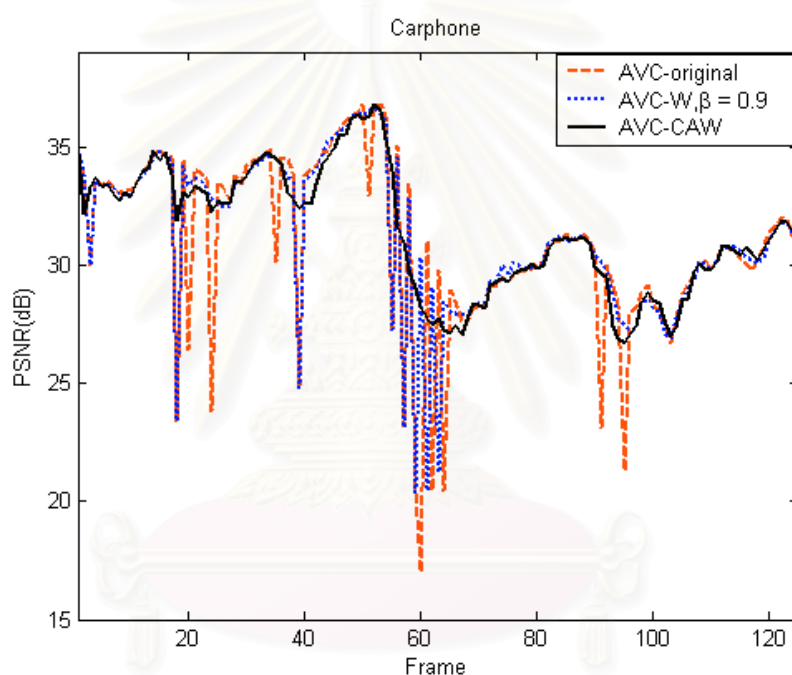
4.2.3 เปรียบเทียบการควบคุมอัตราที่นำเสนอกับการควบคุมอัตราการลงรหัสสัญญาณวีดิทัศน์ขั้นสูง

การควบคุมอัตราที่นำเสนอ คือ การหาค่าถ่วงน้ำหนักบิตเป้าหมายที่เหมาะสม (AVC-W) การกำหนดเงื่อนไขสำหรับการหาค่าควอนไทซ์พารามิเตอร์ (AVC-C) เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายแบบคงที่ (AVC-CW) และ แบบปรับตัวได้ (AVC-CAW)

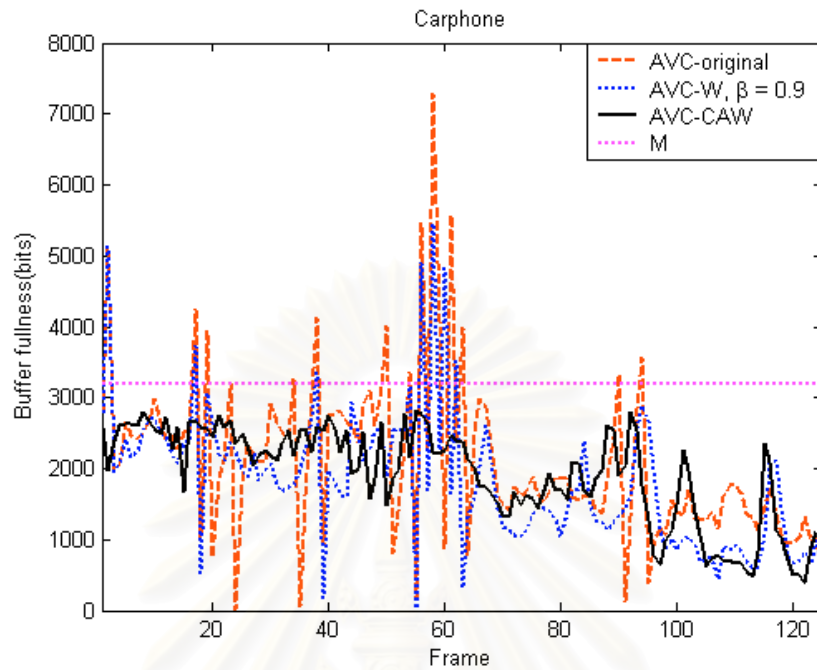
ในหัวข้อนี้เป็นการเปรียบเทียบผลการทดสอบของการควบคุมอัตราที่นำเสนอกับการควบคุมอัตราการลงรหัสสัญญาณวีดิทัศน์ขั้นสูง (AVC) พบว่า ค่าคุณภาพสัญญาณวีดิทัศน์สูงกว่าการควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูง และไม่มีเฟรมกระโดดเกิดขึ้น

4.2.3.1 ผลการทดสอบกับลำดับภาพ Carphone

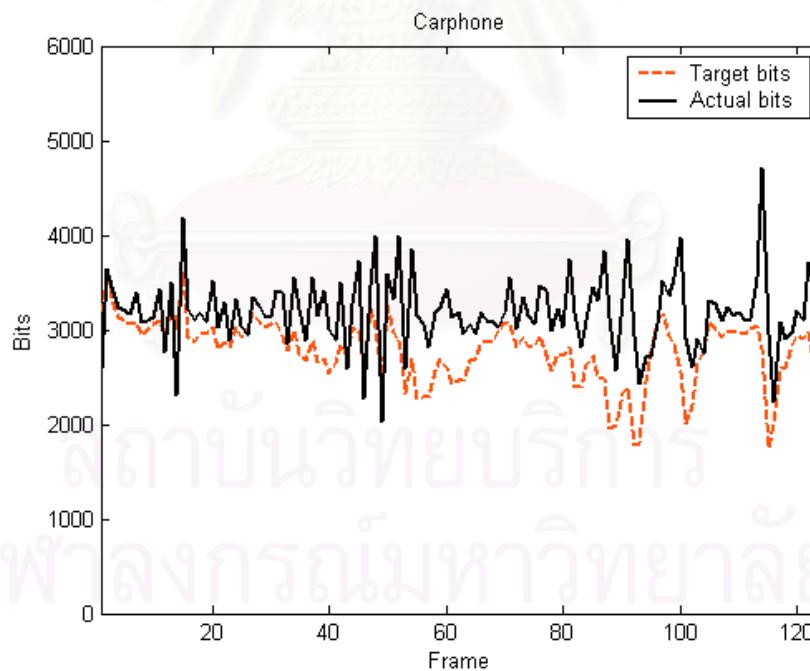
ผลการทดสอบที่ได้แสดงดังรูปที่ 4.25 ถึง 4.27 คือ ค่าคุณภาพสัญญาณวีดิทัศน์ จำนวนบิตสะสมในบัพเฟอร์ และจำนวนบิตเป้าหมายกับจำนวนบิตที่เข้ารหัสจริง ตามลำดับ พบว่า การควบคุมอัตราที่นำเสนอ การหาค่าถ่วงน้ำหนักบิตเป้าหมายที่เหมาะสม (AVC-W) และ การกำหนดเงื่อนไขสำหรับการหาค่าควอนไทซ์ที่เหมาะสม โดยกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายแบบปรับตัวได้ (AVC-CAW) มีค่าคุณภาพสัญญาณวีดิทัศน์ดีกว่าสูงสุดประมาณ 0.57 dB ไม่มีเฟรมกระโดด และจำนวนบิตที่เข้ารหัสกับจำนวนบิตเป้าหมายมีค่าใกล้เคียงกัน



รูปที่ 4.25 ค่าคุณภาพสัญญาณวีดิทัศน์ของลำดับภาพ Carphone



รูปที่ 4.26 จำนวนบิตสะสมในบัฟเฟอร์ของลำดับ Carphone



รูปที่ 4.27 จำนวนบิตเป้าหมายและจำนวนบิตที่เข้ารหัสจริงของลำดับภาพ Carphone

เพื่อให้เห็นถึงประสิทธิภาพของวิธีที่นำเสนอจึงนำลำดับภาพ Carphone แสดงดังรูปที่ 4.28 (ก) ถึง รูปที่ 4.28 (ฉ) โดยรูปที่ 4.28 (ก) ถึง รูปที่ 4.28 (ค) เป็นการเข้ารหัสโดยใช้การควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวิดีโอที่ค่อนข้างสูง และ รูปที่ 4.28 (ง) ถึง รูปที่ 4.28

(ฉ) เป็นการเข้ารหัสโดยใช้การควบคุมอัตราที่นำเสนอ (AVC-CAW) พบว่า เมื่อเข้ารหัสโดยใช้การควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวิดีโอที่สนับสูง เฟรมที่ 20 รูปที่ 4.28 (ข) เกิดการกระโดดข้ามซึ่งทำให้ตัวถอดรหัสต้องแสดงเฟรมที่ 19 ซ้ำ แต่สำหรับการควบคุมอัตราที่นำเสนอสามารถแสดงภาพได้อย่างต่อเนื่อง



(ก) เฟรมที่ 19



(ข) เฟรมที่ 20



(ค) เฟรมที่ 21



(ง) เฟรมที่ 19



(จ) เฟรมที่ 20



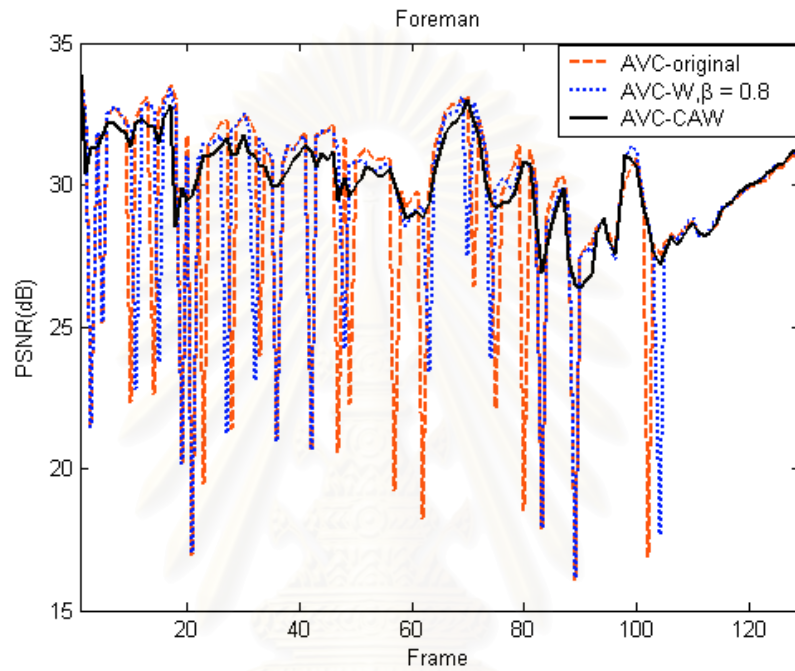
(ฉ) เฟรมที่ 21

รูปที่ 4.28 เปรียบเทียบการควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวิดีโอที่สนับสูงกับการควบคุมอัตราที่นำเสนอของลำดับภาพ Carphone (ก) ถึง (ค) ใช้การควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวิดีโอที่สนับสูง และ (ง) ถึง (ฉ) ใช้การควบคุมอัตราที่นำเสนอ (AVC-CAW)

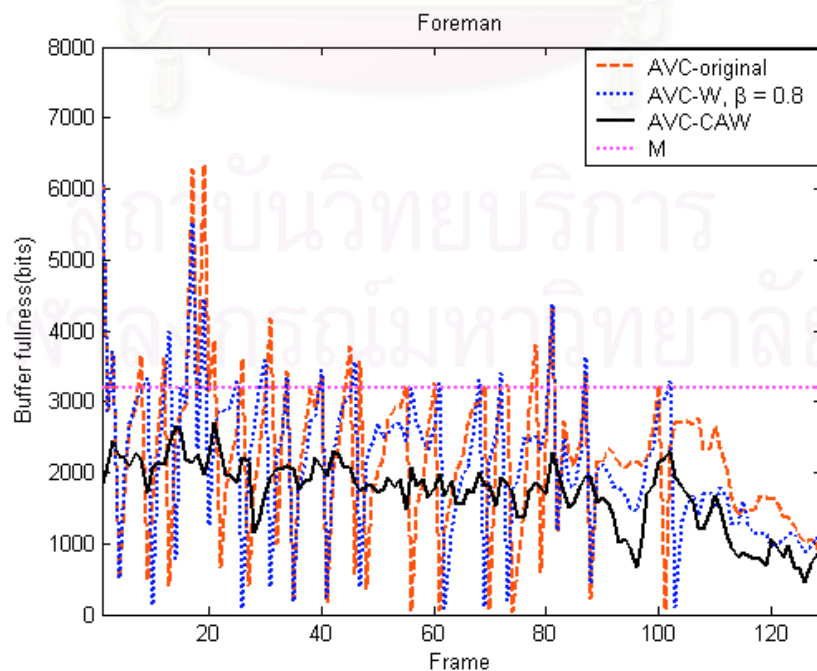
4.2.3.4 ผลการทดสอบกับลำดับภาพ Foreman

ผลการทดสอบที่ได้แสดงดังรูปที่ 4.29 ถึง 4.31 คือ ค่าคุณภาพสัญญาณวิดีโอที่สนับจำนวนบิตสะสมในบัพเฟอร์ และจำนวนบิตเป้าหมายกับจำนวนบิตที่เข้ารหัสจริง ตามลำดับ พบ

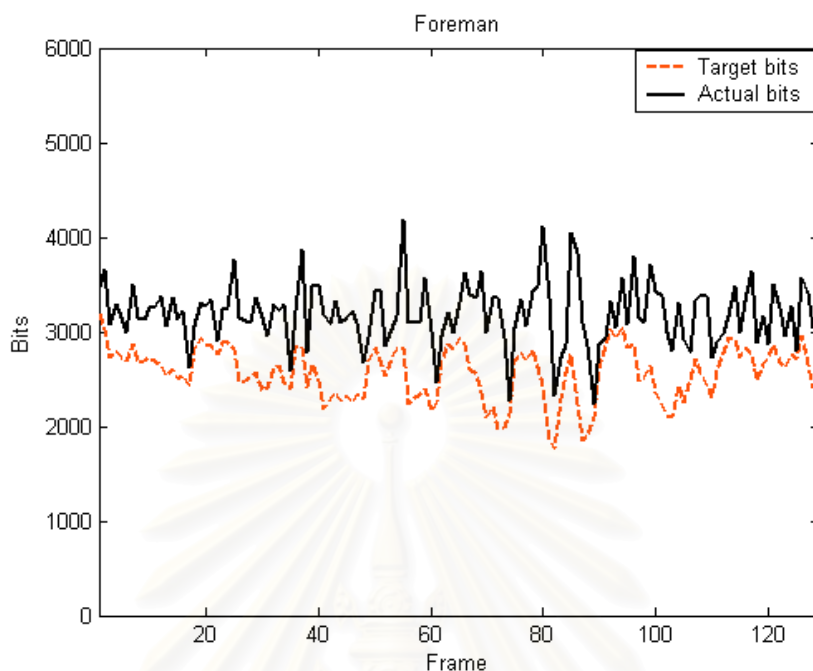
ว่าการควบคุมอัตราที่นำเสนอ การหาค่าถ่วงน้ำหนักบิตเป้าหมายที่เหมาะสม (AVC-W) และ การกำหนดเงื่อนไขสำหรับการหาค่าควอนไทซ์ที่เหมาะสมโดยกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายแบบปรับตัวได้ (AVC-CAW) มีค่าคุณภาพสัญญาณวีดิทัศน์ดีกว่าสูงสุดประมาณ 1.19 dB ไม่มีเฟรมกระโดด และจำนวนบิตที่เข้ารหัสกับจำนวนบิตเป้าหมายมีค่าใกล้เคียงกัน



รูปที่ 4.29 ค่าคุณภาพสัญญาณวีดิทัศน์ของลำดับภาพ Foreman



รูปที่ 4.30 จำนวนบิตสะสมในบัฟเฟอร์ของลำดับ Foreman

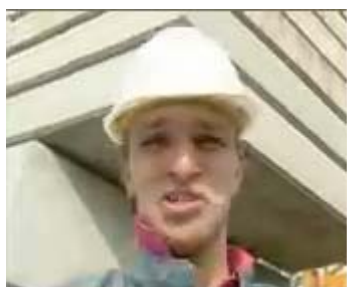


รูปที่ 4.31 จำนวนบิตเป้าหมายและจำนวนบิตที่เข้ารหัสจริงของลำดับภาพ Foreman

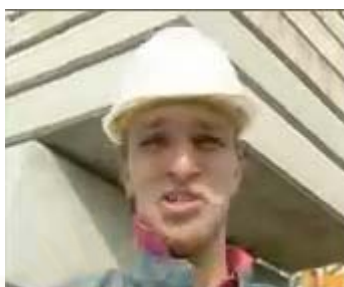
เพื่อให้เห็นถึงประสิทธิภาพของวิธีที่นำเสนอจึงนำลำดับภาพ Foreman แสดงดังรูปที่ 4.32(ก) ถึง รูปที่ 4.32 (ฉ) โดยรูปที่ 4.32 (ก) ถึง รูปที่ 4.32 (ค) เป็นการเข้ารหัสโดยใช้การควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูง และ รูปที่ 4.32 (ง) ถึง รูปที่ 4.32 (ฉ) เป็นการเข้ารหัส โดยใช้การควบคุมอัตราที่นำเสนอ (AVC-CAW) พบว่า เมื่อเข้ารหัสโดยใช้การควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูง เฟรมที่ 52 รูปที่ 4.32 (ข) เกิดการกระโดดข้ามซึ่งทำให้ตัวถอดรหัสต้องแสดงเฟรมที่ 51 ซ้ำ แต่สำหรับการควบคุมอัตราที่นำเสนอสามารถแสดงภาพได้อย่างต่อเนื่อง

4.2.3.4 ผลการทดสอบกับลำดับภาพ Silent

ผลการทดสอบที่ได้แสดงดังรูปที่ 4.33 - 4.35 คือ ค่าคุณภาพสัญญาณวีดิทัศน์ จำนวนบิตสะสมในบัฟเฟอร์ และจำนวนบิตเป้าหมายกับจำนวนบิตที่เข้ารหัสจริง ตามลำดับ พบว่าการควบคุมอัตราที่นำเสนอ การหาค่าถ่วงน้ำหนักบิตเป้าหมายที่เหมาะสม (AVC-W) และ การกำหนดเงื่อนไขสำหรับการหาค่าควอนไทซ์ที่เหมาะสมโดยกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายแบบปรับตัวได้ (AVC-CAW) มีค่าคุณภาพสัญญาณวีดิทัศน์ดีกว่าสูงสุดประมาณ 0.82 dB จำนวน ไม่มีเฟรมกระโดด และจำนวนบิตที่เข้ารหัสกับจำนวนบิตเป้าหมายมีค่าใกล้เคียงกัน



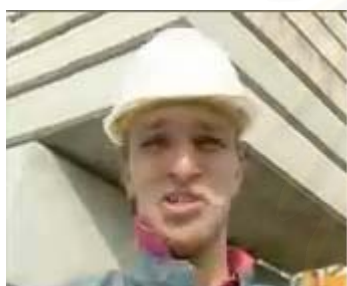
(ก) เฟรมที่ 51



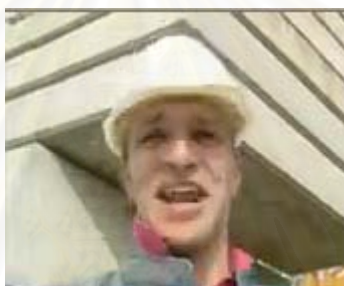
(ข) เฟรมที่ 52



(ค) เฟรมที่ 53



(ง) เฟรมที่ 51



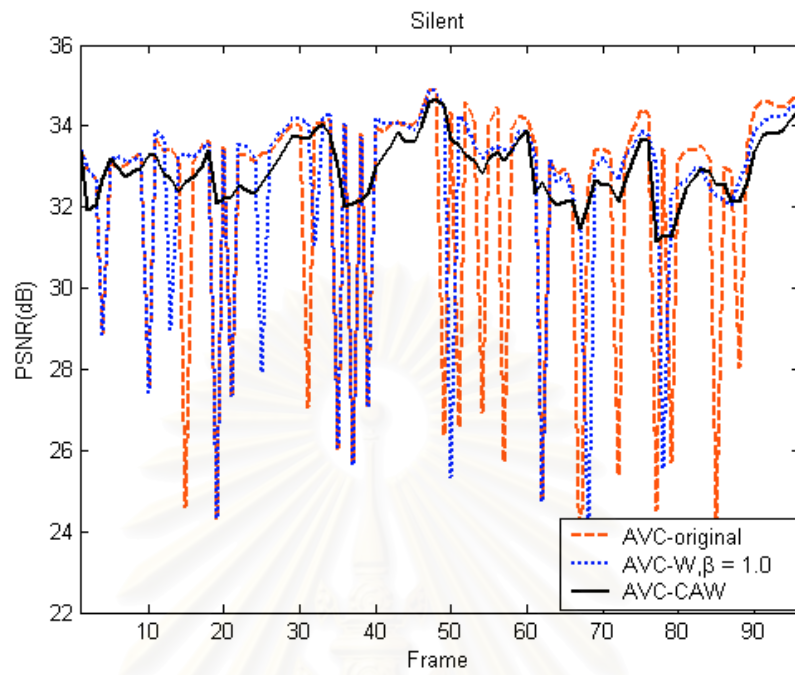
(จ) เฟรมที่ 52



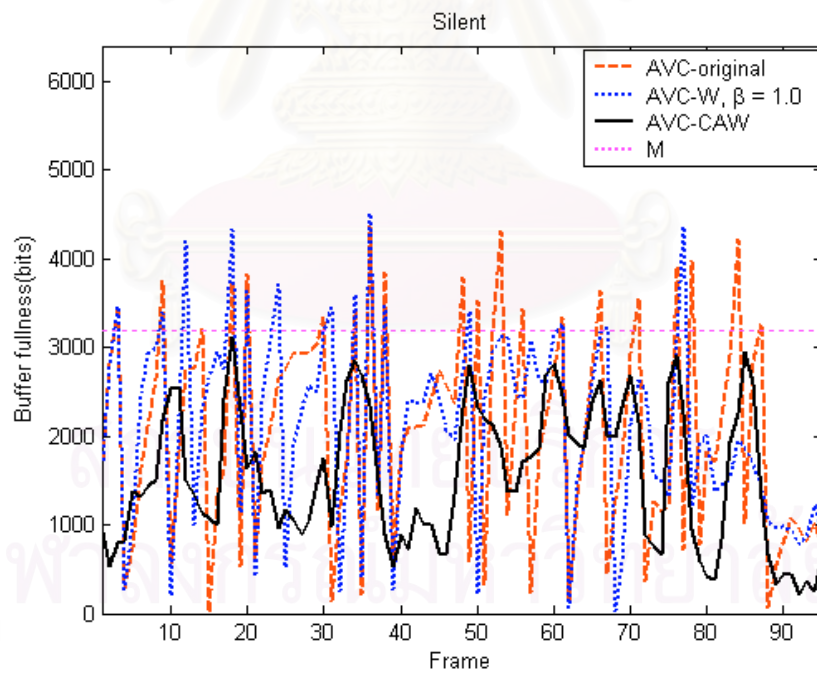
(ฉ) เฟรมที่ 53

รูปที่ 4.32 เปรียบเทียบการควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูงกับการควบคุมอัตราที่นำเสนอของลำดับภาพ Foreman (ก) ถึง (ค) ใช้การควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูง และ (ง) ถึง (ฉ) ใช้การควบคุมอัตราที่นำเสนอ (AVC-CAW)

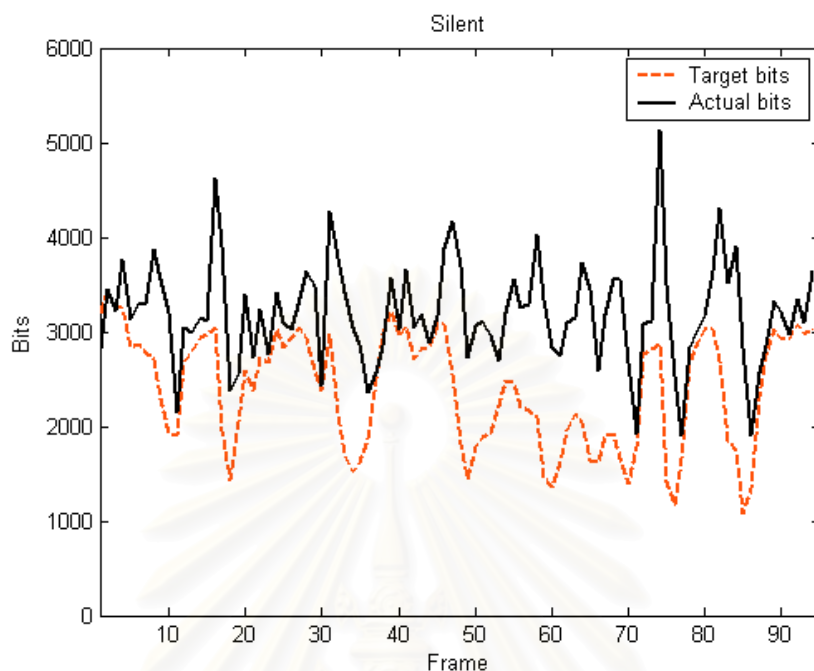
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 4.33 ค่าคุณภาพสัญญาณวิดีโอที่สั้นของลำดับภาพ Silent



รูปที่ 4.34 จำนวนบิตสะสมในบัฟเฟอร์ของลำดับ Silent



รูปที่ 4.35 จำนวนบิตเป้าหมายและจำนวนบิตที่เข้ารหัสจริงของลำดับภาพ Silent

เพื่อให้เห็นถึงประสิทธิภาพของวิธีที่นำเสนอจึงนำลำดับภาพ Silent แสดงดังรูปที่ 4.35 (ก) ถึง รูปที่ 4.36 (จ) โดยรูปที่ 4.36 (ก) ถึง รูปที่ 4.36 (ค) เป็นการเข้ารหัสโดยใช้การควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวิดีโอที่ศน์ชั้นสูง และ รูปที่ 4.36 (ง) ถึง รูปที่ 4.36 (จ) เป็นการเข้ารหัสโดยใช้การควบคุมอัตราที่นำเสนอ (AVC-CAW) พบว่า เมื่อเข้ารหัสโดยใช้การควบคุมอัตราสำหรับ AVC เฟรมที่ 22 รูปที่ 4.36(ข) เกิดการกระโดดข้ามซึ่งทำให้ตัวถอดรหัสต้องแสดงเฟรมที่ 21 ซ้ำ แต่สำหรับการควบคุมอัตราที่นำเสนอสามารถแสดงภาพได้อย่างต่อเนื่อง ซึ่งมีประโยชน์มากสำหรับลำดับภาพภาษามือ เพราะถ้าเกิดเฟรมกระโดด อาจทำให้สื่อสารผิดพลาด

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



(ก) เฟรมที่ 21



(ข) เฟรมที่ 22



(ค) เฟรมที่ 23



(ง) เฟรมที่ 21



(จ) เฟรมที่ 22

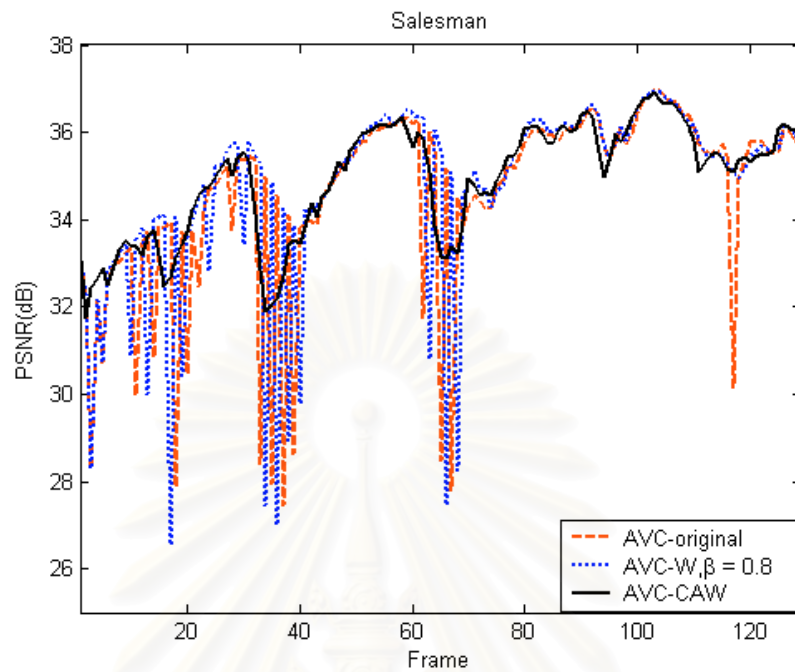


(ฉ) เฟรมที่ 23

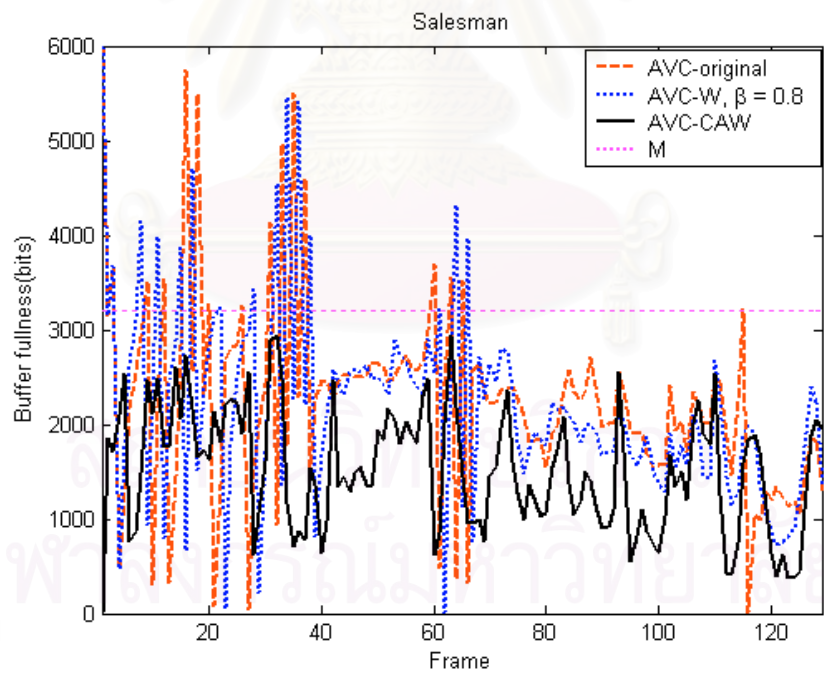
รูปที่ 4.36 เปรียบเทียบการควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูงกับการควบคุมอัตราที่นำเสนอของลำดับภาพ Silent (ก) ถึง (ค) ใช้การควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูง และ (ง) ถึง (ฉ) ใช้การควบคุมอัตราที่นำเสนอ (AVC-CAW)

4.2.3.4 ผลการทดสอบกับลำดับภาพ Salesman

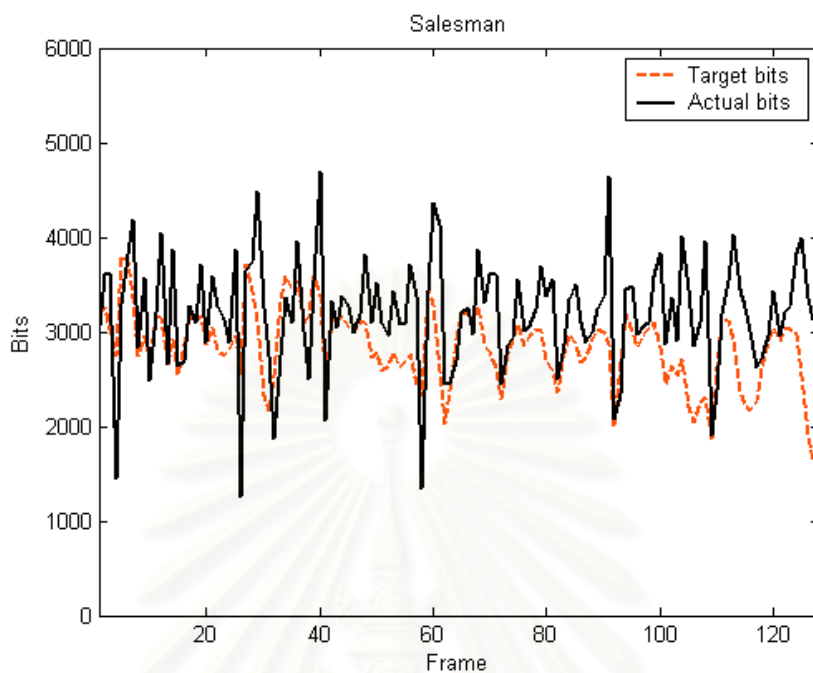
ผลการทดสอบที่ได้แสดงดังรูปที่ 4.37 ถึง 4.39 คือ ค่าคุณภาพสัญญาณวีดิทัศน์จำนวนบิตสะสมในบัพเฟอร์ และจำนวนบิตเป้าหมายกับจำนวนบิตที่เข้ารหัสจริง ตามลำดับ พบว่า การควบคุมอัตราที่นำเสนอ การหาค่าถ่วงน้ำหนักบิตเป้าหมายที่เหมาะสม (AVC-W) และ การกำหนดเงื่อนไขสำหรับการหาค่าควอนไทซ์ที่เหมาะสม โดยกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายแบบปรับตัวได้ (AVC-CAW) มีค่าคุณภาพสัญญาณวีดิทัศน์ดีกว่าสูงสุดประมาณ 0.30 dB ไม่มีเฟรมกระโดด และจำนวนบิตที่เข้ารหัสกับจำนวนบิตเป้าหมายมีค่าใกล้เคียงกัน



รูปที่ 4.37 ค่าคุณภาพสัญญาณวิดีโอที่ต้นของลำดับภาพ Salesman



รูปที่ 4.38 จำนวนบิตสะสมในบัฟเฟอร์ของลำดับ Salesman



รูปที่ 4.39 จำนวนบิตเป้าหมายและจำนวนบิตที่เข้ารหัสจริงของลำดับภาพ Salesman

เพื่อให้เห็นถึงประสิทธิภาพของวิธีที่นำเสนอ จึงนำลำดับภาพ Salesman แสดงดังรูปที่ 4.40 (ก) ถึง รูปที่ 4.40 (ฉ) โดยรูปที่ 4.40 (ก) ถึง รูปที่ 4.40 (ค) เป็นการเข้ารหัสโดยใช้การควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวิดีโอที่ค่อนข้างสูง และ รูปที่ 4.40 (ง) ถึง รูปที่ 4.40 (ฉ) เป็นการเข้ารหัสโดยใช้การควบคุมอัตราที่นำเสนอ (AVC-CAW) พบว่า เมื่อเข้ารหัสโดยใช้การควบคุมอัตราสำหรับ AVC เฟรมที่ 17 รูปที่ 4.40 (ข) เกิดการกระโดดข้ามซึ่งทำให้ตัวถอดรหัสต้องแสดงเฟรมที่ 16 ซ้ำ แต่สำหรับการควบคุมอัตราที่นำเสนอสามารถแสดงภาพได้อย่างต่อเนื่อง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



(ก) เฟรมที่ 16



(ข) เฟรมที่ 17



(ค) เฟรมที่ 18



(ง) เฟรมที่ 16



(จ) เฟรมที่ 17



(ฉ) เฟรมที่ 18

รูปที่ 4.40 เปรียบเทียบการควบคุมอัตราสำหรับมาตรฐานการลงทะเบียนวีดิทัศน์ขั้นสูงกับการควบคุมอัตราที่นำเสนอของลำดับภาพ Salesman (ก) ถึง (ค) ใช้การควบคุมอัตราสำหรับมาตรฐานการลงทะเบียนวีดิทัศน์ขั้นสูง และ (ง) ถึง (ฉ) ใช้การควบคุมอัตราที่นำเสนอ (AVC-CAW)

จากการทดสอบข้างต้นสามารถสรุปได้ดังตารางที่ 4.10 โดยการควบคุมอัตราที่นำเสนอ คือ การกำหนดค่าถ่วงน้ำหนักบิตเป้าหมาย (AVC-W) และ การกำหนดเงื่อนไขสำหรับกำหนดค่าควอนไทซ์พารามิเตอร์เมื่อกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายแบบปรับตัวได้ (AVC-CAW) พบว่า การควบคุมอัตราแบบ AVC-W สามารถลดจำนวนเฟรมกระโดดได้ระหว่าง 1 ถึง 7 เฟรม และเพิ่มค่าคุณภาพสัญญาณวีดิทัศน์สูงสุด 0.40 dB ส่วนการควบคุมอัตราแบบ AVC-CAW ไม่มีเฟรมกระโดด และเพิ่มค่าคุณภาพสัญญาณวีดิทัศน์สูงสุด 1.19 dB

ตารางที่ 4.10 เปรียบเทียบจำนวนเฟรมกระโดดและค่าคุณภาพสัญญาณวีดิทัศน์ ระหว่างการควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวีดิทัศน์ขั้นสูง (AVC) กับ การควบคุมอัตราที่นำเสนอ

ลำดับภาพ	เฟรมที่เข้ารหัส	การควบคุมอัตราสำหรับ AVC		การควบคุมอัตราที่นำเสนอ			
		เฟรมกระโดด	PSNR (dB)	AVC-W		AVC-CAW	
				เฟรมกระโดด	PSNR (dB)	เฟรมกระโดด	PSNR (dB)
Carphone	125	15	30.92	8	31.18 (+0.26)	0	31.49 (+0.57)
Foreman	130	21	29.04	17	29.38 (+0.34)	0	30.23 (+1.19)
Silent	96	20	32.09	14	32.49 (+0.40)	0	33.01 (+0.82)
Salesman	129	16	34.57	15	34.70 (+0.13)	0	34.87 (+0.30)

บทที่ 5

สรุปผลการวิจัย และข้อเสนอแนะ

สรุปผลการวิจัย

วิทยานิพนธ์นี้นำเสนอการควบคุมอัตราสำหรับมาตรฐานการลงรหัสสัญญาณวิทิตทัศน์ชั้นสูง(AVC) เพื่อสามารถสื่อสารที่อัตราบิตต่ำได้อย่างมีประสิทธิภาพ โดยได้ศึกษาผลกระทบของการสะสมบัฟเฟอร์ พบว่ามีจำนวนเฟรมกระโดดมากและลดค่าคุณภาพสัญญาณวิทิตทัศน์ ทั้งนี้เนื่องจากจำนวนบิตเป้าหมายกับจำนวนบิตที่เข้ารหัสจริงแตกต่างกันมาก จึงได้มีการศึกษาพารามิเตอร์ที่มีผลต่อการกำหนดบิตเป้าหมาย ซึ่งพารามิเตอร์ดังกล่าวคือ ค่าถ่วงน้ำหนักบิตเป้าหมาย (β) โดยได้ทำการปรับค่าถ่วงน้ำหนัก โดยวิธีดังกล่าวขอแทนด้วยชื่อย่อว่า AVC-W พบว่า ค่าถ่วงน้ำหนักบิตเป้าหมายมีผลต่อจำนวนเฟรมกระโดดและค่าคุณภาพสัญญาณวิทิตทัศน์ เมื่อปรับค่าให้มามีค่ามากขึ้นจำนวนเฟรมกระโดดลดลงและค่าคุณภาพสัญญาณวิทิตทัศน์ดีขึ้น ดังนั้น การเลือกค่าถ่วงน้ำหนักที่เหมาะสมเป็นสิ่งสำคัญซึ่งค่าถ่วงน้ำหนักบิตเป้าหมายเปลี่ยนแปลงตามลักษณะภาพวิทิตทัศน์จึงได้นำเสนอการปรับค่าถ่วงน้ำหนักแบบปรับตัวได้ โดยจะทำการปรับค่าถ่วงน้ำหนักบิตเป้าหมายทุกเฟรมที่เข้ารหัส

อย่างไรก็ตาม การปรับค่าถ่วงน้ำหนักบิตเป้าหมาย สามารถลดจำนวนเฟรมกระโดดส่วนหนึ่งเท่านั้นเนื่องจากการกำหนดค่าคอนไทซ์พารามิเตอร์มิได้คำนึงถึงจำนวนบิตสะสมในบัฟเฟอร์ในขณะนั้น ดังนั้น จึงนำเสนอเงื่อนไขสำหรับการกำหนดค่าคอนไทซ์พารามิเตอร์โดยคำนึงถึงจำนวนบิตสะสมในบัฟเฟอร์ในขณะนั้นเพื่อป้องกันมิให้มีจำนวนบิตสะสมในบัฟเฟอร์มากกว่าค่าสูงสุดบัฟเฟอร์ โดยวิธีดังกล่าวขอแทนชื่อย่อว่า AVC-C โดยทดสอบรวมกับการกำหนดค่าถ่วงน้ำหนักบิตเป้าหมายแบบคงที่ (AVC-CW) และ แบบปรับตัวได้ (AVC-CAW) โดยผลการทดสอบพบว่า การปรับค่าถ่วงน้ำหนักแบบปรับตัวได้สามารถให้ประสิทธิภาพดีกว่า

เมื่อเปรียบเทียบการควบคุมอัตราแบบ AVC AVC-W และ AVC-CAW พบว่า วิธี AVC-W สามารถลดเฟรมกระโดดอยู่ระหว่าง 1 ถึง 7 เฟรม และเพิ่มค่าคุณภาพสัญญาณวิทิตทัศน์เพิ่มสูงสุด 0.40 dB ส่วนการควบคุมอัตราแบบ AVC-CAW ไม่มีเฟรมกระโดด และเพิ่มค่าคุณภาพสัญญาณวิทิตทัศน์สูงสุด 1.19 dB

ข้อเสนอแนะ

ปัญหาที่เกิดขึ้นในการทำงานวิจัยวิทยานิพนธ์ ได้แก่

เอกสารและข้อมูลที่เกี่ยวข้องกับงานทางด้านการลงรหัสสัญญาณวิทยุยังมีอยู่น้อยมากในประเทศไทย ควรจะมีแผนงานในการจัดหาเอกสารและเข้าร่วมการประชุมสัมมนาขององค์กรต่าง ๆ ที่วิจัยในเรื่องที่เกิดขึ้น

งานที่ควรได้รับการศึกษาหรือพัฒนาต่อไปในอนาคต คือ

1. ปรับปรุงการปรับค่าถ่วงน้ำหนักบิตเป้าหมายให้ออบดีมัล เพื่อนำมาใช้กับโครงสร้างที่นำเสนอ โดยมีจุดประสงค์เพื่อเพิ่มประสิทธิภาพในการเข้ารหัส
2. ปรับปรุงการเข้ารหัสโดยการพิจารณาการเข้ารหัสต้นทางและช่องสัญญาณร่วมกัน (Joint Source Channel Coding) เพื่อนำมาใช้ร่วมกับวิธีนำเสนอเพื่อเพิ่มประสิทธิภาพของการเข้ารหัส
3. พัฒนาการเข้ารหัสโดยนำการแยกส่วนภาพของบริเวณที่สนใจให้มีคุณภาพสัญญาณวิทยุที่ดีขึ้นในบริเวณที่สนใจดีขึ้น

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

รายการอ้างอิง

1. Video Coding of Low Bitrate Communication. ITU-T Recommendation. H.263 1996.
2. Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s. ISO/IEC 11172.
3. Generic coding of moving pictures and associated audio information. ISO/IEC 13818.
4. Information technology-Coding of audio-visual objects. ISO/IEC 14496.
5. J. Ribas-Corbera and S. Lei. Rate Control in DCT Video Coding for Low-Delay Video Communication. IEEE Transactions on Circuit and Systems For Video Technology 9 (February 1999): 172-185.
6. 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.
7. 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.
8. H. Song, J. Kim, and C.-C .J. Kuo. Real-Time encoding frame rate control for H.263+video over internet. Signal Processing: Image Communication 15 (September 1999): 127-148.
9. H. liu and M. El Zarki. Adaptive source rate control for real time wireless video transmission. Mobile Networks and Applications 3 (1998): 49-60.
10. S. Aramvith, I.M. Pao, M.T. Sun. A Rate Control Scheme for Retransmission-based Video Transport over Wireless Channel. Proceeding of the 32nd Asilomar Conference Signals, Systems and Computers (October 1998): 112-116.
11. S. Aramvith and M.T. Sun. A Coding Scheme for Wireless Video Transport with Reduced frame-Skipping. Visual Communication and Image Processing Conference 2000 (June 2000).
12. 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 vol. 13(June 2003): 472-487.
13. Feng Pan, Z. Li, K. Lim, and G. Feng. A Study of MPEG-4 Rate Control Scheme and Its Improvements. IEEE Transactions on Circuit and Systems For Video Technology 13(May 2003): 440-446.
 14. 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.
 15. Siwei Ma, Wen Gao, Peng Gao and Yan Lu. Rate control for advance video coding (AVC) standard. Proceedings of the 2003 International Symposium on Circuits and Systems 2 (May 2003): 892-895.
 16. 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.
 17. JVT-G012r1. Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG Document (March 2003)
 18. 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.
 19. พณชิต กิติปัญญางาม. อัลกอริทึมการเข้ารหัสบล็อกการค้นหาแบบไฮบริดชนิดปรับตัวได้แบบใช้การปรับย้ายจุดศูนย์กลางสำหรับการประมาณการเคลื่อนที่ในการเข้ารหัสสัญญาณวีดิทัศน์. วิทยานิพนธ์ระดับมหาบัณฑิต สาขาวิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2543.
 20. Iain E.G. Richardson. H.264 and MPEG-4 Video Compression. West Sussex, England: Wiley, 2003.
 21. 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.
 22. 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.

23. 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.
24. Yao Wang, Jörn Ostermann and Ya-Qin Zhang. Video Processing and Communication. New Jersey USA: Prentice Hall, 2002.
25. Z. Li, F. Pan, K. Pang Lim, G. Feng, X. Lin and S. Rahardja. JVT-G012r1 software. Signal Processing Program: Institute for InfoComm Research, 2003.
26. C. W. Lin, Y.C. Chen, and M.T. Sun. Dynamic region of interest transcoding for multipoint video conferencing. IEEE Transactions on Circuit and Systems For Video Technology 13 (October 2003): 982-992.
27. Rao, K. R., and Hwang, J. J. Techniques and Standards for Image, Video and Audio Coding. USA :Prentice Hall,1996



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



ภาคผนวก

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทความทางวิชาการที่ได้รับการเผยแพร่

1. Low Delay Rate Control for Advanced Video Coding (AVC) Standard ,Proceeding of Electrical Engineering Conference (EECON26) vol.2 pp. 1052-1057 ,November 2003,Pethburi,Thailand.
2. ได้รับการตอบรับจากงานประชุมทางวิชาการ International Conference on Information Technology (ITCC-2004) สำหรับบทความ Improved Rate Control for Advanced video Coding (AVC) under Low Delay Constraint ซึ่งจะจัดขึ้นในวันที่ 5-7 เมษายน พ.ศ. 2547 ณ เมือง Nevada ประเทศ สหรัฐอเมริกา



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Low Delay Rate Control for Advanced Video Coding (AVC) Standard

Pornthep Navakitkanok and Supavadee Aramvith
Digital Signal Processing Research Laboratory, Department of Electrical Engineering
Chulalongkorn University, Bangkok, Thailand
Tel. 02-218-6909, Fax. 02-218-6912
E-mail: pornthep.n@student.chula.ac.th, supavadee.a@chula.ac.th

Abstract

The upcoming H.264/MPEG-4 Advanced Video Coding (AVC) video compression standard promises a significant improvement over all previous video compression standards. In this paper, we study the impact of low delay constraint on the rate control algorithm proposed for AVC, and show that the performance of AVC rate control is degraded under such delay constraint. This results in PSNR degradation and many skipped frames. We then propose an improved frame-layer rate control scheme to effectively allocate the number of bits and monitor the buffer fullness. Simulation results show that our proposed rate control scheme encodes the video sequences with less frame skipping and with higher PSNR compared to rate control for AVC.

Keywords: Advanced video coding (AVC) standard, H.264/MPEG-4 Part 10, Low delay, Rate control, Frame skipping.

1. Introduction

The high bit rates that result from the various types of multimedia data make their transmission through their intended channels very difficult, especially for digital video. Digital video compression technologies are introduced such that the video can be compressed to transmit at low bit rate medium such as Public Switched Telephone Network (PSTN), Local Area Network (LAN), and wireless networks. These enable the development of several digital video communication applications such as video conferencing.

The upcoming H.264/MPEG-4 AVC video compression standard [1] promises a significant improvement over all previous video compression standards. In terms of coding efficiency, the new standard is expected to provide at least 2x compression improvement over the best previous standards and substantial perceptual quality improvements over both MPEG-2[2] and MPEG-4[3]. The standard, being jointly developed by ITU-T and ISO/IEC, will address the full range of video applications including low bit rate wireless applications, standard-definition and high definition

broadcast television, video streaming over the Internet, delivery of high-definition DVD content, and the highest quality video for digital cinema applications. The ITU-T name for the standard is H.264 (previously called H.26L), while the ISO/IEC name is MPEG-4 Advanced Video Coding (AVC) which will become Part 10 of the MPEG-4 standard.

One important component of video codec is rate control. It is a necessary part of an encoder to allocate the suitable number of bits to each video frame and then smooth out the variable bit rate to constant bit-rate channel. Rate control schemes have been widely studied in standards, like MPEG-2, MPEG-4, H.263[4]. For AVC, there are two different rate control schemes proposed in the literature. One is based on MPEG-2 TM5 rate control [5] and thus is not suitable for low bit rate video applications. Our study bases on another scheme, which is based on fluid traffic model, linear model, and rate distortion optimization (RDO) [6].

In real-time video communications, the end-to-end delay for transmitting video data needs to be very small, particularly in two-way interactive applications such as videophone or video conferencing. If one or several frames occupy higher number of bits than speculated, the additional number of bits will be accumulated in the encoder buffer and resulted in the increase of buffer delay, which is the time needed (at the beginning of a frame's interval) to send the bits remaining from previous frames. When the number of bits in the buffer is too high, the encoder usually skips encoding frames to reduce the buffer delay and avoid buffer overflow. However, the frame skipping produces undesirable motion discontinuity in the encoded video sequence.

In this work, we study the impact of low delay constraint on the rate control algorithm proposed for AVC, and show that the performance of AVC rate control is degraded under such delay constraint. We then propose an improved frame-layer rate control scheme to effectively allocate the number of bits and monitor the buffer fullness. The organization of this paper is as follows. In Section 2, we introduce rate control algorithm for AVC. In Section 3, we illustrate the problem by showing the effects of accumulated encoder buffer of AVC rate control on the video quality under low delay constraint. In Section 4, we

describe the proposed rate-control scheme. Section 5 presents our simulation results and demonstrates that our proposed scheme is effective in reducing the number of frames skipped while improving the average PSNR compared to using AVC. The conclusion is presented in Section 6.

2. Background on Rate Control for AVC

AVC rate control is composed of three layers: GOP layer rate control, frame layer rate control and basic unit layer based on fluid traffic model and linear model.

2.1 GOP layer rate control

In this layer, we need to compute the total number of remaining bits, T_r , for all non-coded frames in each GOP. In the beginning of the i th GOP, the total number of bits allocated for the i th GOP is computed as follows in eq. (1):

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

where N_{gop} denotes the total number of frames in a group of picture (GOP), $n_{i,j}$ ($i=1,2,\dots, j=1,2,\dots, N_{gop}$) denotes the j th frame in the i th GOP, F_r is frame rate, B_s is buffer size, and $B_c(n_{i,j})$ denotes the occupancy of virtual buffer after coding the j th frame computed based on fluid flow traffic model, as defined in eq. (2).

$$B_c(n_{i,j+1}) = \min \left\{ \max \left\{ 0, B_c(n_{i,j}) + A(n_{i,j}) - \frac{u(n_{i,j})}{F_r} \right\}, B_s \right\} \quad (2)$$

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

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

where $A(n_{i,j})$ is the number of bits generated by the j th frame in the i th GOP, $u(n_{i,j})$ is the available channel bandwidth and at j th frame in i th GOP computed as in eq. (3).

$$T_r(n_{i,j}) = T_r(n_{i,j-1}) - A(n_{i,j-1}) \quad (3)$$

2.2 Frame layer rate control

The frame layer rate control scheme consists of two stages: pre-encoding and post-encoding stages.

2.2.1 Pre-encoding stage

The objective of this stage is to compute quantization parameter for all frames via two steps:

Step 1 Determine a target bit for each P frame. Step 1 is composed of the following two sub-steps.

Step 1.1 Macroscopic control (budget allocation among pictures). The bit allocation is implemented by predefining a target buffer level for each P picture as in eq. (4).

$$Tbl(n_{i,j+1}) = Tbl(n_{i,j}) - \frac{B_c(n_{i,2}) - B_s / 8}{N_p - 1} \quad (4)$$

Step 1.2 Microscopic control (target bit rate computation)

The target bit rate, $f(n_{i,j})$, for the j th P frame in the i th GOP is scaled based on the target buffer level, the current buffer level, the frame rate, and the channel bandwidth. It is given as in eq. (5):

$$\tilde{f}(n_{i,j}) = \frac{u(n_{i,j})}{F_r} + \gamma(Tbl(n_{i,j}) - B_c(n_{i,j})) \quad (5)$$

where in γ is a constant and its typical value is 0.75

Further adjustment by a weighted combination of the average number of remaining bits for each frame is given as in eq. (6):

$$f(n_{i,j}) = \beta * \frac{T_r(n_{i,j})}{N_p - j} + (1 - \beta) * \tilde{f}(n_{i,j}) \quad (6)$$

where in β is a constant and its typical value is 0.5

Step 2 Compute the quantization parameter and perform RDO. The details of the step 2 are given in [6].

2.2.2 Post-encoding stage

Update the parameters a_1 and a_2 of linear model, the parameters of quadratic R-D model as in eq. (8).

2.3 Basic Unit Layer Rate Control

A basic unit is defined that a frame is composed of N_{mbpic} macroblocks (MBs). A basic unit is defined to be a group of contiguous MBs which is composed of N_{mbunit} MBs, where N_{mbunit} is a fraction of N_{mbpic} . The total number of basic units in a frame by N_{unit} is computed as in eq. (7).

$$N_{unit} = \frac{N_{mbpic}}{N_{mbunit}} \quad (7)$$

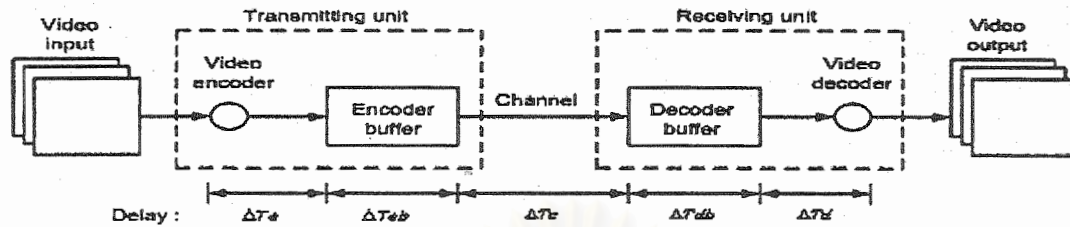


Figure 1. Factors contributing to the end-to-end delay in a video communication system. Adapted from A.Ortega and K.Ramchandran, Rate-distortion methods for image and video compression, *IEEE Signal Processing Magazine* (Nov.1998),15:23-50.Copyright IEEE

The basic unit layer rate control selects the values of quantization parameters of all basic units in a frame, so that the sum of generated bits is close to the frame target $f(n_{i,j})$. The Mean absolute difference (MAD) of current basic unit in the current is predicted base on linear model by that of the basic unit in the co-located position of the previous frame. Suppose that the predicted MAD of current basic unit in the current frame and the actual MAD of basic unit in the co-located position of previous frame are denoted by MAD_{cb} and MAD_{pb} , respectively. The linear prediction model is then given in eq. (8).

$$MAD_{cb} = a_1 * MAD_{pb} + a_2 \quad (8)$$

The details of the basic unit layer rate control are given in [6].

3. Effects of Accumulated Encoder Buffer of AVC Rate Control on Video Quality under Low Delay Constraints

In this section, we demonstrate the impact of the accumulated encoder buffer of AVC rate control on video quality due to low delay constraints.

End-to-end delay from a video source to a destination can be contributed by several factors, as shown in Fig. 1. The end-to-end delay consists of encoder processing delay, ΔT_e (including acquiring the data and encoding); encoder buffer delay, ΔT_{eb} (to smooth the rate variation in the compressed bit stream); transmission delay, ΔT_c (delay caused by transmission itself, which is usually very small, and the delay due to queuing and perhaps retransmission in a packet-based network); decoder buffer delay, T_{db} (to smooth out transmission jitter); and decoder processing delay, ΔT_d (including both decoding and display buffer for constant frame play-out). When the source is pre-compressed, as in video broadcast and streaming applications, the encoding delay is not a critical issue. The ITU-T G.114 standard recommends a

maximum one way delay of 150 ms for acceptable quality of international telephone conversations [7] and is also relevant to interactive video conferencing applications. Thus, in this paper, a delay constraint of 100 ms is assumed.

To show of effects of the accumulated encoder buffer of AVC rate control on video quality, several video sequence were encoded at 32 kbps using AVC rate control.

The bits in the encoder buffer belong to previously encoded frames and must be sent before those of the current frame, thus introducing a delay in the system, denoted as buffer delay. Other sources of delay such as the encoding time can also be important, although the latter is implementation dependent and hence hard to quantify. Nevertheless, we focus on buffer delay since that is generally a major delay component when transmitting real-time video. With our frame skipping, the number of bits in the buffer $B_c(n_{i,j})$ will not be

larger than M (at the beginning of the current frame's interval), and hence the buffer delay will be less than

$\frac{M}{u(n_{i,j})}$ seconds. By default, we set $M = \frac{u(n_{i,j})}{F_r}$ (as

in [10]), and hence the maximum buffer delay (τ) is defined as in eq. (9).

$$\tau = \frac{M}{u(n_{i,j})} = \frac{1}{F_r} \text{ seconds} \quad (9)$$

i.e., bit rate is 32 kbps, frame rate is 10 fps then maximum buffer delay is 100 ms

Figs. 2(a) and 2(c) illustrate the buffer fullness level. Figs. 2(b) and (d) show PSNR of video sequence. On the average, the PSNR on AVC rate control under low delay constraint is dropped around 0.74 and 1.9 dB for "Claire" and "Foreman" sequences, respectively. The number of frames skipped are 18 and 23 frames, for "Claire" and "Foreman" sequences, respectively.

4. Proposed Low Delay Rate Control for AVC

The simulation results from Section 3 state that using AVC rate control under low delay constraints results in the encoder buffer fill-up and many skipped frames. In this section, we propose rate control scheme to improve the video quality.

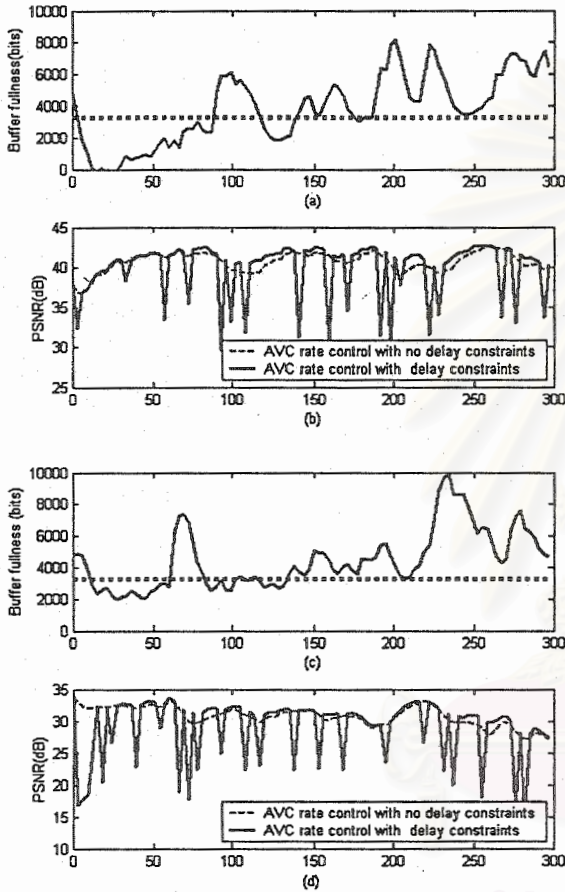


Figure 2. Simulation using the video sequence coded with AVC rate control at 32 kbps and a target frame-rate of 10 fps (a)-(b) Using "Claire" video sequence (c)-(d) Using "Foreman" video sequence. (a),(c) Buffer fullness level (in bits) at each frame interval, maximum delay constraints(dashed line). (b),(d) PSNR comparison between AVC rate control (solid line) with delay constraints and with no delay constraint (dashed line).

4.1 Frame skipping

Before the current frame is (inter) coded, we update the number of bits in the encoder buffer ($B_c(n_{i,j})$). It is given as in eq. 10.

$$B_c(n_{i,j}) = B_c(n_{i,j-1}) + (A(n_{i,j-1}) - \frac{u(n_{i,j-1})}{F_r}) \quad (10)$$

The initial buffer fullness is set to zero. The updated number of bits in the buffer equals the previous buffer fullness plus the bits used for the previous frame and minus the number of bits taken by the channel per frame interval (i.e. 3200 bits). If buffer fullness is larger than or equal to encoder buffer size, the encoder skips encoding frames until the buffer fullness below encoder buffer size, as shown in eq. (11).

$$\begin{aligned} & \text{while } B_c(n_{i,j}) > M \\ & \{ B_c(n_{i,j}) = B_c(n_{i,j}) - \frac{u(n_{i,j-1})}{F_r} \} \quad (11) \end{aligned}$$

4.2 GOP layer rate control

In low delay applications, the typical format of a GOP used is IPPPPP..., and there is one GOP in a sequence. The total number of remaining bits T_r for all non-coded frames is computed accordingly as in eq. (12).

$$T_r(n_{1,0}) = \frac{u(n_{1,1})}{F_r} * N_{gop} \quad (12)$$

The buffer occupancy is computed as in eq. (13).

$$\begin{aligned} B_c(n_{1,j+1}) &= \max\{0, B_c(n_{1,j}) + A(n_{1,j}) - \frac{u(n_{1,j})}{F_r}\} \\ B_c(n_{1,1}) &= 0 \end{aligned} \quad (13)$$

The total number of remaining bits T_r for the j th frame is computed accordingly as in eq. (14).

$$T_r(n_{1,j}) = T_r(n_{1,j-1}) - A(n_{1,j-1}) \quad (14)$$

4.3 Frame layer rate control

The frame layer rate control scheme consists of two stages: pre-encoding and post-encoding stages.

4.3.1 Pre-encoding stage.

The objective of this stage is to compute quantization parameter for all frames. If buffer fullness is greater than 25% of maximum value, there is a tendency that buffer will overflow. Thus, the weighting parameters in eqs. (5) and (6) will be set as $\gamma = 1.0, \beta = 0.1$. The algorithm can be summarized as in eq. (15).

$$\text{If } \{ B_c(n_{1,j}) > kM \} \quad (15)$$

$$\tilde{f}(n_{1,j}) = \frac{u(n_{1,j})}{F_r} + \gamma(Tbl(n_{1,j}) - B_c(n_{1,j}))$$

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

where k is constant

4.3.2 Post-encoding stage

The buffer fullness will be checked again for the chance of frame skipping. If buffer fullness is larger than or equal to encoder buffer size, the encoder skips encoding frames until the buffer fullness level falls below encoder buffer size. We use buffer constraint condition to protect buffer overflow, thus reduce frame skipping.

5. Simulation Results

We perform simulations with three test video sequences: Foreman, Claire, and Carphone with following parameter settings: QCIF(4:2:0) format, an input frame rate of 30 fps, an encoded frame rate of 10 fps, a QP for the first I picture = 32, a QP for the first P picture = 34, a Bit rate of 32kbps, a buffer size of 3200 bits (corresponds to 100 ms delay), and other parameter settings are shown in Table 1. The PSNR calculation when frame skipping occurs is done according to the rate-control tests in MPEG-4, which stated: it was decided that when a frame is skipped, the previous encoded frame should be used in the PSNR computation because the decoder displays the previous encoded frame in place of the skipped one. Simulation results, compared between AVC rate control and our proposed scheme, are shown in Fig. 3 for the "Claire" sequence and Fig. 4 for the "Foreman" sequence. With the proposed frame-layer rate-control, we achieve a better PSNR (up to 0.75 dB improvement) with up to 50% reduced number of frames skipped, as shown in Table 2.

TABLE 1
PARAMETER CONDITION

MV resolution	¼ pel
Hadamard	on
RD optimization	off
Search Range	±16(QCIF)
Reference frames	1
Restrict Search Range	2
Symbol mode	UVLC
GOP structure	IPP
Frametobeencoded	100

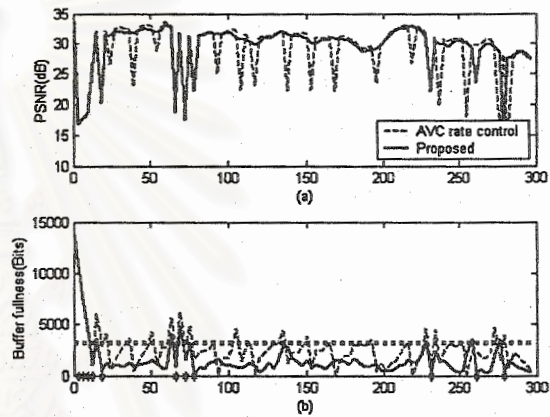


Fig. 3. Simulations using the "Claire" video sequence coded with AVC (Rate control) at 32kbps and a target frame rate of 10 fps. (a) PSNR comparison between AVC and our proposed frame-layer rate-control scheme. (b) Buffer fullness (in bits) comparison at each frame interval. The asterisk at the bottom denotes a skipped frame of our proposed scheme.

TABLE 2

COMPARISON OF NUMBER OF FRAMES SKIPPED AND AVERAGE PSNR FOR AVC AND OUR PROPOSED SCHEME

Video Sequence	Total Frames	AVC Rate Control		Our Proposed Rate control		
		# Frames Skipped	PSNR (dB)	# Frames Skipped	% of Frames Skipped Reduction	PSNR (dB)
"Claire"	100	18	39.85	9	50%	40.20 (+0.25)
"Foreman"	100	23	28.91	11	52%	29.66 (+0.75)
"Carphone"	100	22	30.94	15	31%	31.28 (+0.34)

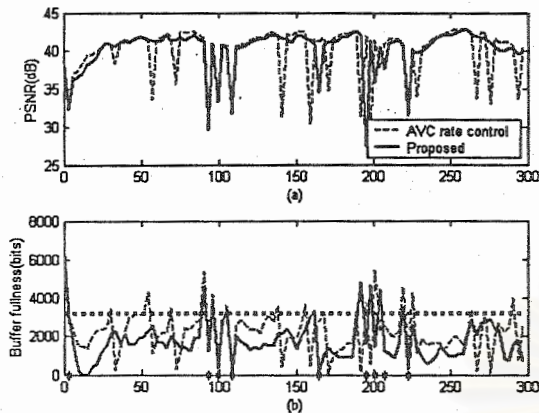


Fig. 4. Simulations using the "Foreman" video sequence coded with AVC (Rate control) at 32kbps and a target frame rate of 10 fps. (a) PSNR comparison between AVC and our proposed frame-layer rate-control scheme. (b) Buffer fullness (in bits) comparison at each frame interval. The asterisk at the bottom denotes a skipped frame of our proposed scheme.

6. Conclusions

In this paper, we consider the scenario of low delay scenario for two-way video communication. We show that, the performance of AVC rate control under low delay constraint is dropped as there are many skipped frames and PSNR degradation. To improve the overall quality, we proposed rate control scheme for real time communication, which take into the effects of the encoder buffer fill-up. As a result, our proposed rate-control schemes encode the video sequences with better motion continuity and with better PSNR quality.

7. Acknowledgement

This research is in part supported by Research Unit for Center of Excellence in Digital Signal Processing, Ratchadaphisek Somphot Endowment, Chulalongkorn University.

Reference

- [1] Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG Document JVT-G050r1, 2003-06-02
- [2] ITU-T and ISO/IEC JTC1, "Generic coding of moving pictures and associated audio information - Part 2: Video," ITU-T Recommendation H.262 - ISO/IEC 13818-2 (MPEG-2), Nov. 1994.
- [3] ISO/IEC JTC1, "Coding of audio-visual objects - Part 2: Visual," ISO/IEC 14496-2 (MPEG-4 visual version 1), Apr. 1999; Amendment 1 (version 2), Feb. 2000; Amendment 4 (streaming profile), Jan. 2001.
- [4] ITU-T, "Video coding for low bit rate communication," ITU-T Recommendation H.263; version 1, Nov. 1995; version 2, Jan. 1998; version 3, Nov. 2000..
- [5] Siwei Ma, Wen Gao, Peng Gao and Yan Lu, "Rate control for Advanced video coding (AVC) standard," *IEEE Circuits and Systems, 2003. ISCAS '03. Proceedings of the 2003 International Symposium on*, vol. 2, pp. 892-895, May 25-28, 2003

- [6] Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG Document JVT-G012r1, 2003-03-10
- [7] Yao Wang, Jörn Ostermann, Ya-Qin Zhang, "Video Processing and Communication," Prentice Hall, New Jersey, USA, 2002
- [8] J. Ribas-Corbera and S. Lei, "Rate Control in DCT Video Coding for Low-Delay Video Communications," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, pp. 172-185, Feb. 1999.
- [9] Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG Software JVT-G012r1 software, 2002-10-04
- [10] J. Ribas-Corbera and S. Lei, "Rate Control in DCT Video Coding for Low-Delay Video Communications," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, pp. 172-185, Feb. 1999.



Pornthep Navakitkanok received the B.Eng. degree in electrical engineering from Kasetsart University, Bangkok, Thailand, in 2001. He is working toward the M.Eng. degree at Chulalongkorn University, Bangkok, Thailand. His research interests include rate-control and rate-distortion optimization for advanced video coding (AVC) standard.



Supavadee Aramvith received the B.S. (first class honors) degree in computer science from Mahidol University, Bangkok, Thailand, in 1993. She received the M.S. and Ph.D. degrees in electrical engineering from the University of Washington, Seattle, USA, in 1996 and 2001, respectively. She is currently with Department of Electrical Engineering, Chulalongkorn University, Bangkok, Thailand. Her research interests include rate-control, low bit-rate video coding, video transmission over wireless and heterogeneous networks, and development of multimedia communication applications such as video conferencing system and E-learning based courseware.

Improved Rate Control for Advanced Video Coding (AVC) Standard under Low Delay Constraint

Pornthep Navakitkanok and Supavadee Aramvith

*Digital Signal Processing Research Laboratory, Department of Electrical Engineering
Chulalongkorn University, Bangkok, Thailand*

Tel. 02-218-6909, Fax. 02-218-6912

E-mail: pornthep.n@student.chula.ac.th, supavadee.a@chula.ac.th

Abstract

The upcoming H.264/MPEG-4 Advanced Video Coding (AVC) video compression standard promises a significant improvement over all previous video compression standards. In this paper, we study the impact of low delay constraint on the rate control algorithm proposed for AVC, and show that the performance of AVC rate control is degraded under such low delay constraint. This results in PSNR degradation and many skipped frames. We then propose an improved rate control scheme to effectively allocate the number of bits and monitor the buffer fullness. Simulation results show that our proposed rate control scheme encodes the video sequences with less frame skipping and with higher PSNR compared to rate control for AVC.

1. Introduction

The upcoming H.264/MPEG-4 AVC video compression standard [1] promises a significant improvement over all previous video compression standards. In terms of coding efficiency, the new standard is expected to provide at least 2x compression improvement over the best previous standards and substantial perceptual quality improvements over both MPEG-2[2] and MPEG-4[3]. The standard, being jointly developed by ITU-T and ISO/IEC, will address the full range of video applications including low bit rate wireless applications, standard-definition and high definition broadcast television, video streaming over the Internet, delivery of high-definition DVD content, and the highest quality video for digital cinema applications. The ITU-T name for the standard is H.264, while the ISO/IEC name is MPEG-4 Advanced Video Coding (AVC) which will become Part 10 of the MPEG-4 standard.

One important component of video codec is rate control. It is a necessary part of an encoder to allocate the suitable number of bits to each video frame and then smooth out the variable bit rate to constant bit-rate channel. Rate control schemes have been widely studied in standards, like MPEG-2, MPEG-4, H.263[4]. For AVC,

there are two different rate control schemes proposed in the literature. One is based on MPEG-2 TM5 rate control [5] and thus is not suitable for low bit rate video applications. Our study bases on another scheme, which is based on fluid traffic model, linear model, and rate distortion optimization (RDO) [6].

In real-time video communications, the end-to-end delay for transmitting video data needs to be very small, particularly in two-way interactive applications such as videophone or video conferencing. If one or several frames occupy higher number of bits than speculated, the additional number of bits will be accumulated in the encoder buffer and resulted in the increase of buffer delay, which is the time needed (at the beginning of a frame's interval) to send the bits remaining from previous frames. When the number of bits in the buffer is too high, the encoder usually skips encoding frames to reduce the buffer delay and avoid buffer overflow. However, the frame skipping produces undesirable motion discontinuity in the encoded video sequence.

In this work, we study the impact of low delay constraint on the rate control algorithm proposed for AVC, and show that the performance of AVC rate control is degraded under such low delay constraint. We then propose an improved rate control scheme to effectively allocate the number of bits and monitor the buffer fullness. The organization of this paper is as follows. In Section 2, we introduce rate control algorithm for AVC. In Section 3, we illustrate the problem by showing the effects of accumulated encoder buffer of AVC rate control on the video quality under low delay constraint. In Section 4, we describe the proposed rate-control scheme. Section 5 presents our simulation results and demonstrates that our proposed scheme is effective in reducing the number of frames skipped while improving the average PSNR compared to using AVC. The conclusion is presented in Section 6.

2. Background on Rate Control for AVC [6]

AVC rate control is composed of three layers: GOP layer rate control, frame layer rate control and basic unit

layer based on fluid traffic model and linear model.

2.1 GOP layer rate control

In this layer, we need to compute the total number of remaining bits, $T_r(n_{i,j})$, $n_{i,j}$ denotes the j th frame in the i th GOP for all non-coded frames in each GOP. The details of the GOP layer rate control are given in [6].

2.2 Frame layer rate control

The frame layer rate control scheme consists of two stages: pre-encoding and post-encoding stages.

2.2.1 Pre-encoding stage. The objective of this stage is to compute quantization parameter for all frames via two steps:

Step 1 Determine a target bit for each P frame. Step 1 is composed of the following two sub-steps.

Step 1.1 Macroscopic control (budget allocation among pictures). The bit allocation is implemented by predefining a target buffer level for each P picture, as shown in eq. (1).

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

,where N_p denotes number of P frame in this GOP, B_s is buffer size, and $B_c(n_{i,j})$ denotes the occupancy of virtual buffer after coding the j th frame

Step 1.2 Microscopic control (target bit rate computation). The target bit rate, $f(n_{i,j})$, for the j th P frame in the i th GOP is scaled based on the target buffer level, current buffer level, frame rate F_r , and channel bandwidth $u(n_{i,j})$. It is given in eq. (2).

$$\tilde{f}(n_{i,j}) = \frac{u(n_{i,j})}{F_r} + \gamma(T_{bl}(n_{i,j}) - B_c(n_{i,j})) \quad (2)$$

,where frame rate, γ is a constant and its default value is 0.75. Further adjustment by a weighted combination of the average number of remaining bits for each frame is given, as shown in eq. (3).

$$f(n_{i,j}) = \beta * \frac{T_r(n_{i,j})}{N_p - j} + (1 - \beta) * \tilde{f}(n_{i,j}) \quad (3)$$

,where in β is a constant and its default value is 0.5

Step 2 Compute the quantization parameter and perform RDO. The quantization parameter \hat{Q}_j corresponding to the target bit is then computed by using the modified quadratic model as in eq. (4) [10].

$$\frac{f(n_{i,j}) - H(n_{i,j})}{M(n_{i,j})} = a_1 \times \hat{Q}_j^{-1} + a_2 \times \hat{Q}_j^{-2} \quad (4)$$

,where H denotes the bits used for header, motion vectors, M_j is MAD current frame, Q_j is the quantization level used for the current frame, a_1 and b_1 denote first and the second-order coefficients. To maintain the smoothness of visual quality among successive frames, the quantization parameter is further adjusted and then used to perform RDO and details of RDO are given in [6].

2.2.2 Post-encoding stage. Update the parameters a_1 and a_2 of linear model as in eq. (5) and the parameters of quadratic R-D model.

2.3 Basic Unit Layer Rate Control

A basic unit is defined that a frame is composed of N_{mbpic} macroblocks (MBs). A basic unit is defined to be a group of contiguous MBs which is composed of N_{mbunit} MBs, where N_{mbunit} is a fraction of N_{mbpic} . The total number of basic units in a frame by N_{unit} is denoted by $N_{unit} = N_{mbpic} / N_{mbunit}$

The objective of basic unit layers is computed the quantization parameter for all basic unit in a frame and sum of generated bits close to frame target bits $f(n_{i,j})$ via six steps:

Step 1 Compute the number of texture bits $T_{j,k}$ for the current basic unit, details of step 1 are given in [6]

Step 2 Predict the MAD of current basic unit in the current frame and actual MAD of basic unit in co-located position of previous frame are denoted by MAD_{cb} and MAD_{pb} , respectively the linear prediction model is then given in eq. (5).

$$MAD_{cb} = a_1 \times MAD_{pb} + a_2 \quad (5)$$

Step 3 Compute the quantization parameter of current basic unit Q_{cb} by using the quadratic R-D model. We need to consider the following three cases:

Case 1 For the first basic unit in the current frame, the quantization parameter is set to the average value of quantization parameters for all basic units in the previous frame, Q_{apf} , denoted as $Q_{cb} = Q_{apf}$

Case 2 If the number of bits left to encode the rest of basic unit in a frame is less than zero, the quantization parameter should be set to be greater than the quantization parameter of the previous basic unit, Q_{pb} , such that the sum of generated bits is close to

$f(n_{i,j})$, as shown in eq. (6).

$$\hat{Q}_{cb} = Q_{pb} + DQuant \quad (6)$$

To maintain the smoothness of visual quality, the quantization parameter is further bounded, as shown in eq. (7).

$$Q_{cb} = \max \{ Q_{apf} - \Delta, \min \{ 1, Q_{apf} + \Delta, \hat{Q}_{cb} \} \} \quad (7)$$

If N_{mbunit} is less than the total number of MBs in a row, the value of Δ is set to 3 and otherwise Δ is set to 6.

Case 3 If the number of bits left to encode the rest of basic unit in a frame is greater than zero, a quantization parameter Q_{cb} is computed by using the quadratic model. The value is bounded in order to reduce the blocking artifacts. The detail are in [6]

Step 4 Perform RDO for all MBs in the current basic unit.

Step 5 Update the number of remaining bits and the number of non-coded basic units for the current frame.

Step 6 After coding a whole frame, Q_{apf} is updated.

3. Effects of Accumulated Encoder Buffer of AVC Rate Control on Video Quality under Low Delay Constraints

To show of effects of the accumulated encoder buffer of AVC rate control on video quality, several video sequence were encoded at 32 kbps using AVC rate control.

The bits in the encoder buffer belong to previously encoded frames and must be sent before those of the current frame, thus introducing a delay in the system, denoted as buffer delay. Other sources of delay such as the encoding time can also be important, although the latter is implementation dependent and hence hard to quantify. Nevertheless, we focus on buffer delay since that is generally a major delay component when transmitting real-time video. To impose a low delay constraint, the number of bits in the buffer $B_c(n_{i,j})$ must not be greater than M (at the beginning of the current frame's interval), otherwise the frame will be skipped. Hence the buffer delay will be less than

$\frac{M}{u(n_{i,j})}$ seconds. By default, we set $M = \frac{u(n_{i,j})}{F_r}$, as in

[9], and hence the maximum buffer delay (τ) is

$\frac{M}{u(n_{i,j})} = \frac{1}{F_r}$ seconds. In our study, bit rate of 32 kbps,

and frame rate of 10 fps which correspond to the maximum buffer delay of 100 ms are used.

Figs. 1(a) illustrates the buffer fullness level. When the buffer fullness level reaches a certain level, i.e., 3200 bits, the frame is skipped. The frame skipping results in discontinued motion. In simulations, we take the first 300

frame of video sequence which has originally a frame rate of 30 fps. It is then encoded with a target frame rate of 10 fps, i.e., 96 coded video frames for "Foreman" sequence, in which 20 video frames are skipped under low delay constraint. AVC rate control, when operated under low delay constraint, also results in the PSNR degradations especially the sharp drops due to the frame skipplings, i.e., the PSNR on AVC rate control under low delay constraint is dropped around 1.52 dB for "Foreman" sequence, as shown in fig. 1(b).

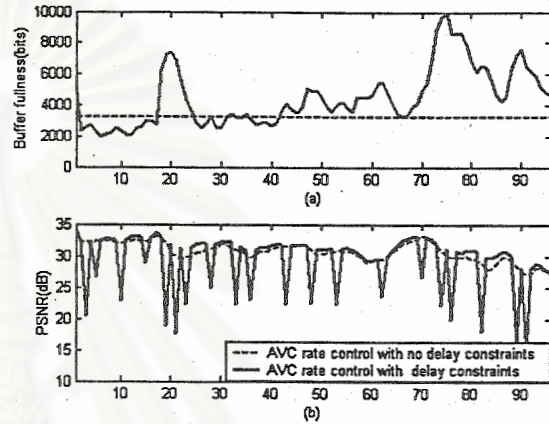


Figure 1. Simulation using the video sequence coded with AVC rate control at 32 kbps and a target frame-rate of 10 fps (a)-(b) Using "Foreman video sequence. (a) Buffer fullness level (in bits) at each frame interval, maximum delay constraints (dashed line). (b) PSNR comparison between AVC rate control (solid line) with delay constraints and with no delay constraint (dashed line)"

4. Proposed Low Delay Rate Control for AVC

The simulation results from Section 3 state that using AVC rate control under low delay constraints results in the encoder buffer fill-up and many skipped frames. In this section, we propose rate control scheme to improve the video quality.

4.1 Frame skipping

Before the current frame is (inter) coded, we update the number of bits in the encoder buffer ($B_c(n_{i,j})$), as in eq. (8).

$$B_c(n_{i,j}) = B_c(n_{i,j-1}) + (A(n_{i,j-1}) - (\frac{u(n_{i,j-1})}{F_r})) \quad (8)$$

The initial buffer fullness is set to zero. The updated number of bits in the buffer equals the previous buffer fullness plus the bits used for the previous frame and minus the number of bits taken by the channel per frame interval (i.e. 3200 bits). If buffer fullness is larger than or

equal to encoder buffer size, the encoder skips encoding frames until the buffer fullness below encoder buffer size, as shown in eq. (9).

$$\begin{aligned} & \text{while } B_c(n_{i,j}) > M \\ & \{ B_c(n_{i,j}) = B_c(n_{i,j}) - \frac{u(n_{i,j-1})}{F_r} \} \end{aligned} \quad (9)$$

4.2 GOP layer rate control

In low delay applications, the typical format of a GOP used is IPPPP...., and there is one GOP in a sequence. The total number of remaining bits T_r for all non-coded frames is computed accordingly as in eq. (10).

$$T_r(n_{1,0}) = \frac{u(n_{1,1})}{F_r} * N_{gop} \quad (10)$$

The buffer occupancy is computed as in eq. (11).

$$\begin{aligned} B_c(n_{1,j+1}) &= \max\{0, B_c(n_{1,j}) + A(n_{1,j}) - \frac{u(n_{1,j})}{F_r}\} \\ B_c(n_{1,1}) &= 0 \end{aligned} \quad (11)$$

The total number of remaining bits T_r for the j th frames is computed accordingly as in eq. (12).

$$T_r(n_{1,j}) = T_r(n_{1,j-1}) - A(n_{1,j-1}) \quad (12)$$

4.3 Frame layer rate control

The frame layer rate control scheme consists of two stages: pre-encoding and post-encoding stages.

4.3.1 Pre-encoding stage. The objective of this stage is to compute quantization parameter for all frames. From our study, we observe that there is a tendency buffer overflow if the buffer fullness is greater than 75% level. Thus, if the buffer is greater than 75% level, the weighting parameters used in eq. (2) and (3) will be set as $\gamma = 1.0, \beta = 0.1$, otherwise the default value of γ and β , as stated in section 2.1.1, will be used.

To get a better target bit-rate estimation, we need to consider complexity each frame from MAD. If the MAD of previous frame is greater than average MAD and buffer fullness less than 75% level then increase the target bit rate by 10%. The target bit-rate estimation can be further adjusted, as shown in eq. (13).

$$\begin{aligned} & \text{if } \{(MAD_{j-1} > MAD_{Average}) \text{ and } B_c(n_{1,j}) < nM\} \\ & f(n_{1,j}) = 1.1(f(n_{1,j})) \end{aligned} \quad (13)$$

,where MAD_{j-1} is the MAD previous frame, n is constant

$MAD_{Average}$ is the average of MAD from the previous three frames.

4.3.2 Post-encoding stage. The buffer fullness will be checked again for frame skipping. If buffer fullness is

larger than or equal to encoder buffer size, the encoder skips encoding frames until the buffer fullness level falls below encoder buffer size. We use buffer constraint condition to protect buffer overflow, thus reduce frame skipping.

4.4 Basic Unit Layer Rate Control

In basic unit layer rate control, the computation of quantization parameter for the first basic unit (case 1) and when there are enough bits to encode the rest of the basic unit in a frame (case 3) is done according to the original algorithm as stated in section 2.3. For case 2 where there is not enough bits to encode the rest of the basic unit, the quantization parameter, \hat{Q}_{cb} , will be set to be greater than that of previous basic unit, i.e., by setting the value of $DQUANT$ to 1. However, the final quantization parameter, Q_{cb} , will be bounded by less than three levels of average value of quantization parameters in previous frame, as stated in eq. (6). As a result, we observe that the sum of generated bits for all basic unit is always much higher than the expected frame target bit rate. Thus, there is higher possibility for excess encoder buffer accumulation and frame skipped, as we already show in the simulation in fig. 1. To alleviate this problem, we then propose a condition for selecting the quantization parameter by monitoring buffer fullness level at the basic unit layer. We define the current buffer fullness after encoding the basic unit k , B_k , as shown in eq. (14).

$$B_k = B_c(n_{1,j}) + \sum_{i=1}^k B_{b-i}, \quad 0 \leq B_k \leq 2M \quad (14)$$

where B_{b-i} is the number of bits used to encode basic unit i . If B_k is lower than threshold $h\%$ level, the final quantization parameter is set as in eq. (7), otherwise we retain the value calculated in eq. (6). In the simulation, h is set to 1.75. The algorithm can be summarized as in eq. (15).

$$\begin{aligned} & \text{if } (f(n_{1,j}) < 0) \\ & \hat{Q}_{cb} = Q_{pb} + DQUANT \end{aligned} \quad (15)$$

$$\begin{aligned} & \text{if } (B_k < hM) \{ \\ & Q_{cb} = \max\{1, Q_{apf} - \Delta, \min\{51, Q_{apf} + \Delta, \hat{Q}_{cb}\}\} \\ & \text{else} \end{aligned}$$

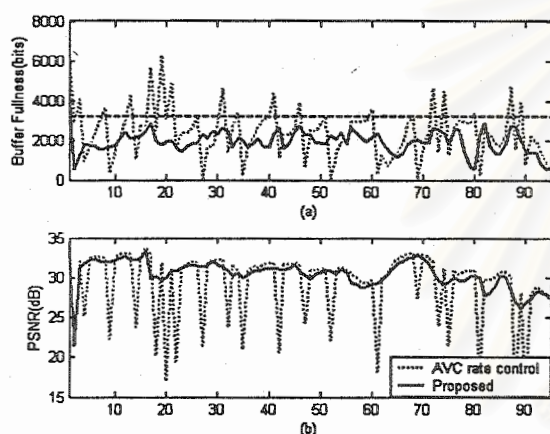
$$Q_{cb} = \hat{Q}_{cb}$$

5. Simulation Results

We perform simulations by using JVT reference software [8] with three test video sequences: Foreman, Claire, and Carphone with following parameter settings:

QCIF(4:2:0) format, an input frame rate of 30 fps, an encoded frame rate of 10 fps, a QP for the first I picture = 32, a QP for the first P picture = 34, a Bit rate of 32kbps, a buffer size of 3200 bits (corresponds to 100 ms delay). The PSNR calculation when frame skipping occurs is done according to the rate-control tests in MPEG-4 [11], which stated that it was decided that when a frame is skipped, the previous encoded frame should be used in the PSNR computation because the decoder displays the previous encoded frame in place of the skipped one.

The comparison between AVC rate control and our proposed scheme is shown in Fig. 2 for the "Foreman" sequence. With the proposed rate-control, we achieve a better PSNR (up to 1.27 dB improvement) with up to 100% reduced number of frames skipped, as shown in Table 1.



"Figure 2. Simulations using the "Foreman" video sequence coded with AVC (Rate control) at 32kbps and a target frame rate of 10 fps. (a) PSNR comparison between AVC and our proposed rate-control scheme. (b) Buffer fullness (in bits) comparison at each frame interval"

6. Conclusions

In this paper, we consider the scenario of low delay scenario for two-way video communication. We show that, the performance of AVC rate control under low delay constraint is dropped as there are many skipped

frames and PSNR degradation. To improve the overall quality, we proposed rate control scheme for real time communication, which take into the effects of the encoder buffer fill-up. As a result, our proposed rate-control schemes encode the video sequences with better motion continuity and with better PSNR quality

7. Acknowledgement

This research is in part supported by Research Unit for Center of Excellence in Digital Signal Processing, Ratchadaphisek Somphot Endowment, Chulalongkorn University.

Reference

- [1] Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG Document JVT-G050r1, June.2003
- [2] ITU-T and ISO/IEC JTC1, "Generic coding of moving pictures and associated audio information - Part 2: Video," ITU-T Recommendation H.262 - ISO/IEC 13818-2 (MPEG-2), Nov. 1994.
- [3] ISO/IEC JTC1, "Coding of audio-visual objects - Part 2: Visual," ISO/IEC 14496-2 (MPEG-4 visual version 1), Apr. 1999; Amendment 1 (version 2), Feb. 2000; Amendment 4 (streaming profile), Jan. 2001.
- [4] ITU-T, "Video coding for low bit rate communication," ITU-T Recommendation H.263; version 1, Nov. 1995; version 2, Jan. 1998; version 3, Nov. 2000.
- [5] Siwei Ma, Wen Gao, Peng Gao and Yan Lu, "Rate control for Advanced video coding (AVC) standard," *IEEE Circuits and Systems, 2003. ISCAS '03. Proceedings of the 2003 International Symposium on*, vol. 2, May 25-28, 2003, pp. 892-895.
- [6] Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG Document JVT-G012r1, Mar.2003.
- [7] Yao Wang, Jörn Ostermann, Ya-Qin Zhang, *Video Processing and Communication*, Prentice Hall, New Jersey, USA, 2002
- [8] Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG Software JVT-G012r1 software, Mar.2003.
- [9] J. Ribas-Corbera and S. Lei, "Rate Control in DCT Video Coding for Low-Delay Video Communications," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, pp. 172-185, Feb. 1999.
- [10] H.J.Lee and T.H. Chiang and Y.Q.Zhang, "Scalable Rate Control for MPEG-4 Video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, Sep. 2000, pp. 878-894.
- [11] Video Group, "Text of ISO/IEC 14496-2 MPEG-4 Video VM - version 8.0," ISO/IEC JTC1/SC29/WG11 Coding of Moving Picture and Associated Audio MPEG 1796, Stockholm, Sweden, July 1997.

"Table 1. Comparison of number of frame skipped and average PSNR for and our proposed scheme"

Video Sequence	Total Frames	AVC Rate Control		Our Proposed Rate Control		
		# Frames Skipped	PSNR (dB)	# Frames Skipped	% of Frames Skipped Reduction	PSNR (dB)
"Claire"	99	17	39.70	0	100%	40.38 (+0.68)
"Foreman"	96	20	29.16	1	95%	30.43 (+1.27)
"Carphone"	97	20	31.13	2	90%	31.75 (+0.62)

ประวัติผู้เขียนวิทยานิพนธ์

นายพรเทพ นวกิจกนก เข้ารับการศึกษาในหลักสูตรวิศวกรรมศาสตรบัณฑิต คณะ
วิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ ในปีการศึกษา 2540 สำเร็จการศึกษาปริญญาตรี
วิศวกรรมศาสตรบัณฑิต สาขาวิชาวิศวกรรมไฟฟ้า ในปีการศึกษา 2543 และเข้าศึกษาต่อในหลักสูตร
วิศวกรรมศาสตรมหาบัณฑิต ในสังกัดห้องปฏิบัติการกรรมวิธีสัญญาณดิจิทัล ภาควิชาวิศวกรรมไฟฟ้า
ที่จุฬาลงกรณ์มหาวิทยาลัยในปีการศึกษา 2544



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย