

การปรับปรุงคุณภาพสำหรับวีดิทัศน์ภาษามือ H.264 ผ่านการส่งแบบไร้สายอัตราบิตต่ำ



บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)
เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ ที่ส่งผ่านทางบัณฑิตวิทยาลัย

The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository (CUIR)
are the thesis authors' files submitted through the University Graduate School.

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมไฟฟ้า ภาควิชาวิศวกรรมไฟฟ้า
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2557
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

QUALITY IMPROVEMENT FOR H.264 SIGN LANGUAGE VIDEO OVER LOW BIT-RATE
WIRELESS TRANSMISSION

Mr. Sirawich Subongkot



A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Engineering Program in Electrical Engineering
Department of Electrical Engineering
Faculty of Engineering
Chulalongkorn University
Academic Year 2014
Copyright of Chulalongkorn University

ศิริวิชญ์ สุขงกช : การปรับปรุงคุณภาพสำหรับวีดิทัศน์ภาษามือ H.264 ผ่านการส่งแบบไร้สายอัตราบิตต่ำ (QUALITY IMPROVEMENT FOR H.264 SIGN LANGUAGE VIDEO OVER LOW BIT-RATE WIRELESS TRANSMISSION) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: ผศ. ดร.สุภาวดี อร่ามวิทย์, 86 หน้า.

สำหรับคนหูหนวก ภาษามือถูกใช้เป็นวิธีหลักในการติดต่อสื่อสารแทนการสื่อสารด้วยวาจา ด้วยความก้าวหน้าของเทคโนโลยีจึงเป็นเรื่องธรรมดาที่จะเห็นบริการโทรศัพท์วีดิทัศน์หรือการประชุมทางวีดิทัศน์ที่คนหูหนวกสามารถใช้บริการเหล่านี้สำหรับการติดต่อสื่อสารโดยใช้การแสดงท่าภาษามือระหว่างกัน อย่างไรก็ตามวีดิทัศน์ที่มีการส่งผ่านโครงข่ายทั้งแบบมีสายหรือไร้สายที่มีความกว้างของช่องสัญญาณต่ำและง่ายต่อการเกิดข้อผิดพลาด การเสื่อมของคุณภาพเป็นสิ่งที่หลีกเลี่ยงไม่ได้ ภาพที่บกพร่องนี้ทำให้การแปลความหมายที่ถูกต้องแทบจะเป็นไปไม่ได้เลย

ในวิทยานิพนธ์นี้ ผลงานซึ่งประกอบด้วยวิธีการสามแบบถูกนำเสนอด้วยจุดมุ่งหมายเพื่อสร้างวีดิทัศน์ที่ต้านทานต่อความผิดพลาด ลดผลกระทบของการแพร่กระจายความผิดพลาดและเพื่อนำเสนอวิธีการจัดสรรบิตที่เหมาะสมสำหรับวีดิทัศน์ภาษามือ วิธีในลำดับแรกแผนการคืนสภาพมาโครบล็อกภายในถูกนำเสนอร่วมกับแบบจำลองทางภาษา ข้อมูลป้อนกลับและบิตเป้าหมายเพื่อคำนวณจำนวนมาโครบล็อกภายในที่เหมาะสมต่อเฟรม การคืนสภาพมาโครบล็อกภายในถูกจัดสรรในบริเวณใบหน้า มือและลำตัว ลำดับที่สองการจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้ถูกใช้ในการจัดสรรมาโครบล็อกเข้าสู่กลุ่มสไลซ์ที่แตกต่างกันตามแบบจำลองทางภาษาเพื่อลดความน่าจะเป็นที่มาโครบล็อกที่สำคัญจะเกิดการสูญหาย ในลำดับที่สามการปรับปรุงการควบคุมอัตราบิตในระดับเฟรมและมาโครบล็อกถูกนำเสนอเพื่อการจัดสรรบิตที่ชาญฉลาดเพื่อเพิ่มความชัดเจนในการแปลความหมายของวีดิทัศน์ภาษามือที่ถูกเข้ารหัส

ผลการทดลองถูกประเมินคุณภาพในเชิงปริวิสัยและความชัดเจนในการแปลความหมาย ผลการทดลองแสดงให้เห็นว่าวิธีที่นำเสนอสามารถปรับปรุงคุณภาพของวีดิทัศน์ภาษามือที่ถูกบีบอัดได้อย่างมีนัยสำคัญและสามารถเพิ่มความชัดเจนในการแปลความหมายในการสนทนาวีดิทัศน์ภาษามือได้เป็นอย่างดี

ภาควิชา วิศวกรรมไฟฟ้า

ลายมือชื่อนิสิต

สาขาวิชา วิศวกรรมไฟฟ้า

ลายมือชื่อ อ.ที่ปรึกษาหลัก

ปีการศึกษา 2557

5370674121 : MAJOR ELECTRICAL ENGINEERING

KEYWORDS: H.264/AVC / FMO / INTRA MACROBLOCK REFRESH / ERROR RESILIENTR / SIGN LANGUAGE

SIRAWICH SUBONGKOT: QUALITY IMPROVEMENT FOR H.264 SIGN LANGUAGE VIDEO OVER LOW BIT-RATE WIRELESS TRANSMISSION. ADVISOR: ASST. PROF. SUPAVADEE ARAMVITH, Ph.D., 86 pp.

For the Deaf people, Sign language is used as a primary mode of their communication, in place of verbal communication. With technology advancement, it is common to see video phone and video conference services where deaf can utilize these services for their signing communication. However as video has to transmit over low bandwidth and error prone wired or wireless network, the quality degradation is inevitably expected. With impaired images, the accurate interpretation is not possible.

In this thesis, three contribution are presented with the aim to make the video more resilient to error, to reduced effects of error propagation, and to propose suitable bit allocation method for signing video. Firstly, an intra macroblock refresh scheme is proposed. By incorporating linguistic model, feedback information, and target bits, the optimal number of macroblocks per frame is computed. The intra refresh macroblock are assigned in face, hand, and body region. Secondly, flexible macroblock ordering is used to by assigning macroblocks to different slice groups based on linguistic model to reduce the probability that important macroblock will be lost. Lastly, improved frame layer and macroblock layer rate control are proposed to intelligently allocate bit to increase intelligibility of coded signing videos.

The experimental results are evaluated in terms of objective quality and intelligibility and showed that the proposed method can significantly improve compressed signing video quality and can increase intelligibility in signing video conversation as well.

Department: Electrical Engineering

Student's Signature

Field of Study: Electrical Engineering

Advisor's Signature

Academic Year: 2014

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลงได้ด้วยคำแนะนำและความช่วยเหลือจากอาจารย์ที่ปรึกษา วิทยานิพนธ์คือ ผู้ช่วยศาสตราจารย์ ดร.สุภาวดี อร่ามวิทย์ ที่ให้คำแนะนำ เป็นที่ปรึกษา และแนะแนวทางในการแก้ปัญหาที่เกิดขึ้นระหว่างการทำวิทยานิพนธ์ ตลอดจนให้โอกาสในการเรียนรู้และสัมผัสงานวิจัยต่างๆหลากหลายด้านซึ่งถือว่าเป็นโอกาสที่ดีในการเรียนรู้สิ่งใหม่

ขอขอบพระคุณบิดาและมารดา ซึ่งให้การสนับสนุนด้านการศึกษา ค่าใช้จ่ายและให้กำลังใจในระหว่างการศึกษาระดับมหาบัณฑิต

ขอขอบคุณรุ่นพี่ เพื่อนและน้องๆ ในภาควิชาวิศวกรรมไฟฟ้าที่ช่วยเหลือในเรื่องต่างๆ ทั้งด้านการเรียนและกิจกรรมต่างๆ ที่ทำให้การเรียนระดับมหาบัณฑิตมีความสุขสนุกสนานและมีสีสันกว่าที่คิด

ขอขอบคุณบัณฑิตวิทยาลัยสำหรับ “ทุนอุดหนุนวิทยานิพนธ์สำหรับนิสิต” ที่ช่วยสนับสนุนค่าใช้จ่ายบางส่วนในการทำการวิทยานิพนธ์ครั้งนี้

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญตาราง.....	ฅ
สารบัญภาพ	ฉ
บทที่ 1 บทนำ	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของวิทยานิพนธ์	3
1.3 ขอบเขตของวิทยานิพนธ์.....	3
1.4 ประโยชน์ที่คาดว่าจะได้รับ.....	4
1.5 ขั้นตอนและวิธีการดำเนินงาน	4
1.6 ลำดับของการนำเสนอเนื้อหา.....	4
บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง	6
2.1 มาตรฐานการเข้ารหัสสัญญาณวีดิทัศน์ H.264 [8]	6
2.1.2 ชนิดของรูปภาพ	7
2.1.2 การเข้ารหัสภายในเฟรม.....	8
2.1.3 การเข้ารหัสระหว่างเฟรม.....	11
2.1.4 การแปลงและการแปลงผกผัน	13
2.1.5 การแจกหน่วยและการกราดค่าสัมประสิทธิ์.....	14
2.1.6 การเข้ารหัสเอนโทรปี.....	15
2.1.7 การควบคุมอัตราสำหรับการเข้ารหัสสัญญาณวีดิทัศน์ขั้นสูง.....	15

2.1.7.1 การควบคุมอัตราระดับกลุ่มภาพ (GOP level rate control).....	16
2.1.7.2 การควบคุมอัตราระดับเฟรม (Frame level rate control).....	16
2.1.7.3 การควบคุมอัตราระดับหน่วยพื้นฐานย่อย (Basic unit level rate control).....	20
2.1.8 อัตราบิด – ความเพี้ยนที่เหมาะสม (Rate-distortion Optimization)	22
2.2 เทคนิคการต้านทานความผิดพลาดในมาตรฐานการเข้ารหัส H.264/AVC	23
2.2.1 ลำดับกลุ่มสไลซ์แบบกำหนดเอง (Arbitrary Slice Order: ASO).....	23
2.2.2 กลุ่มสไลซ์ซ้ำซ้อน (Redundant Slices).....	23
2.2.3 การตัดแบ่งข้อมูล (Data Partitioning)	23
2.2.4 การคืนสภาพภายใน (Intra Refreshing).....	23
2.2.5 การจัดเรียงมาโครบล็อกแบบยืดหยุ่น (Flexible Macroblock Ordering: FMO).....	24
2.3 ตัวจำลองช่องสัญญาณไร้สาย	25
2.3.1 ลักษณะเฉพาะของช่องสัญญาณไร้สาย	26
2.3.2 รูปแบบของมาโครบล็อกที่ถอดรหัสไม่ได้	28
2.3.2.1 มาโครบล็อกที่ถอดรหัสไม่ได้ชนิดที่ 1	28
2.3.2.1 มาโครบล็อกที่ถอดรหัสไม่ได้ชนิดที่ 2	28
2.4 แบบจำลองทางภาษามือ	29
2.5 การวัดคุณภาพสัญญาณวีดิทัศน์	29
2.5.1 ค่าอัตราส่วนสัญญาณยอดต่อสัญญาณรบกวน (Peak Signal to Noise Ratio: PSNR).....	30
2.5.2 ค่าความชัดเจนในการแปลความหมายของวีดิทัศน์ภาษามือ (Computation Intelligibility model: CIM) [5]	30
2.5.3 ค่าสัมประสิทธิ์การตัดสินใจ (Coefficient of Determination : R^2).....	30
2.6 เอกสารและงานวิจัยที่เกี่ยวข้อง.....	31

บทที่ 3 เทคนิคที่นำเสนอ	35
3.1 แผนการคืนสภาพมาโครบล็อกภายใน (Intra macroblock refresh scheme)	35
3.1.1 แบบจำลองอัตราคืนสภาพมาโครบล็อกภายใน	35
3.1.2 การจัดสรรการคืนสภาพมาโครบล็อกภายในบริเวณใบหน้าและมือของผู้ทำภาษา มือ.....	37
3.1.3 การจัดสรรการคืนสภาพมาโครบล็อกภายในบริเวณลำตัวของผู้ทำภาษามือ	38
3.2 การจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้ (Flexible Macroblock Ordering: FMO)	39
3.3 การปรับปรุงการควบคุมอัตราระดับเฟรมและมาโครบล็อกสำหรับมาตรฐานการเข้ารหัส H.264/AVC.....	41
3.3.1 การปรับปรุงการควบคุมอัตราระดับเฟรมสำหรับมาตรฐานการเข้ารหัส H.264/AVC	41
3.3.2 การปรับปรุงการควบคุมอัตราระดับหน่วยพื้นฐานย่อยสำหรับมาตรฐานการเข้ารหัส H.264/AVC	46
บทที่ 4 ผลการจำลอง.....	49
4.1 พารามิเตอร์ที่ใช้ในการจำลอง	49
4.1.1 พารามิเตอร์ของช่องสัญญาณไร้สายจำลอง	49
4.1.2 พารามิเตอร์ของตัวเข้ารหัส/ตัวถอดรหัสวีดิทัศน์ H.264.....	50
4.1.2.1 ตัวเข้ารหัส.....	50
4.1.2.2 ตัวถอดรหัส	50
4.2 การวัดประสิทธิภาพของเทคนิคที่ได้นำเสนอ	51
4.2.1 ค่าคุณภาพวีดิทัศน์ PSNR.....	51
4.2.2 ค่าความชัดเจนในการแปลความหมายของวีดิทัศน์ภาษามือ CIM.....	51
4.3 ผลการจำลอง.....	51
4.3.1 การหาอัตราการคืนสภาพมาโครบล็อกภายในที่เหมาะสม.....	51

4.3.2 การจัดการสารเคมีในสภาพมาโครบล็อกภายในบริเวณใบหน้าและมือของผู้ทำท่าภาษามือ.....	54
4.3.3 การจัดการสารเคมีในสภาพมาโครบล็อกภายในบริเวณลำตัวของผู้ทำท่าภาษามือ	57
4.3.4 การใช้การจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้.....	61
4.3.5 การปรับปรุงการควบคุมอัตราระดับเฟรมสำหรับมาตรฐานการเข้ารหัส H.264/AVC	68
4.3.5.1 แบบจำลองบิตส่วนหัว.....	68
4.3.5.2 การวัดความซับซ้อนภายในเฟรม.....	69
4.3.6 การปรับปรุงการควบคุมอัตราระดับหน่วยพื้นฐานย่อยสำหรับมาตรฐานการเข้ารหัส H.264/AVC	72
บทที่ 5 สรุปผลการจำลองและข้อเสนอแนะ	78
5.1 สรุปผลการจำลอง.....	78
5.2 ข้อเสนอแนะสำหรับงานวิจัยในอนาคต	80
รายการอ้างอิง	81
ประวัติผู้เขียนวิทยานิพนธ์	86

สารบัญตาราง

ตารางที่ 4.1	พารามิเตอร์ที่ใช้ในซอฟต์แวร์ช่องสัญญาณจำลอง	49
ตารางที่ 4.2	ข้อกำหนดในการทดสอบหาอัตราคืนสภาพภายในที่เหมาะสม	52
ตารางที่ 4.3	เปรียบเทียบค่าเฉลี่ย PSNR บริเวณใบหน้าและมือของผู้ทำท่าภาษามือกรณีเฟตดิ้ง ช้า	55
ตารางที่ 4.4	เปรียบเทียบค่าเฉลี่ย PSNR บริเวณใบหน้าและมือของผู้ทำท่าภาษามือกรณีเฟตดิ้ง เร็ว	55
ตารางที่ 4.5	เปรียบเทียบค่าเฉลี่ย PSNR บริเวณใบหน้า มือและลำตัวกรณีเฟตดิ้งช้า	58
ตารางที่ 4.6	เปรียบเทียบค่าเฉลี่ย PSNR บริเวณใบหน้า มือและลำตัวกรณีเฟตดิ้งเร็ว	58
ตารางที่ 4.7	เปรียบเทียบค่าเบี่ยงเบนมาตรฐานของค่าความเพี้ยนที่กรณีเฟตดิ้งช้า	62
ตารางที่ 4.8	เปรียบเทียบค่าเบี่ยงเบนมาตรฐานของค่าความเพี้ยนที่กรณีเฟตดิ้งเร็ว	62
ตารางที่ 4.9	เปรียบเทียบค่าเฉลี่ย CIM ในบริเวณของผู้ทำท่าภาษามือกรณีเฟตดิ้งช้า	63
ตารางที่ 4.10	เปรียบเทียบค่าเฉลี่ย CIM ในบริเวณของผู้ทำท่าภาษามือกรณีเฟตดิ้งเร็ว	64
ตารางที่ 4.11	ค่า R^2 ที่อัตราบิต 64 K	68
ตารางที่ 4.12	ค่า R^2 ที่อัตราบิต 96 K	68
ตารางที่ 4.13	ค่า R^2 ที่อัตราบิต 128 K	69
ตารางที่ 4.14	ค่า R^2 ที่อัตราบิต 160 K	69
ตารางที่ 4.15	เปรียบเทียบค่าเฉลี่ย CIM ในบริเวณของผู้ทำท่าภาษามือกรณีเฟตดิ้งช้า	71
ตารางที่ 4.16	เปรียบเทียบค่าเฉลี่ย CIM ในบริเวณของผู้ทำท่าภาษามือกรณีเฟตดิ้งเร็ว	71
ตารางที่ 4.17	เปรียบเทียบค่าเฉลี่ย CIM ในบริเวณของผู้ทำท่าภาษามือกรณีเฟตดิ้งช้า	72
ตารางที่ 4.18	เปรียบเทียบค่าเฉลี่ย CIM ในบริเวณของผู้ทำท่าภาษามือกรณีเฟตดิ้งเร็ว	73

สารบัญภาพ

รูปที่ 2.1	บล็อกไดอะแกรมการเข้ารหัสวีดิทัศน์ H.264.....	6
รูปที่ 2.2	ความสัมพันธ์ระหว่างเฟรมแบบ I P และ B ในการทำนายการเคลื่อนที่.....	8
รูปที่ 2.3	ตำแหน่งส่วนประกอบความส่องสว่างขนาด 4X4 จุดภาพ	9
รูปที่ 2.4	รูปแบบการเข้ารหัสภายในเฟรมของส่วนประกอบความส่องสว่างขนาด 4X4 จุดภาพ.....	10
รูปที่ 2.5	รูปแบบการเข้ารหัสภายในเฟรมของส่วนประกอบความส่องสว่างขนาด 16X16 จุดภาพ..	11
รูปที่ 2.6	แถวบน : การแบ่งบล็อกขนาด 16x16 จุดภาพ, แถวล่าง : การแบ่งบล็อกขนาด 8x8 จุดภาพ.....	12
รูปที่ 2.7	การประมาณการเคลื่อนที่	12
รูปที่ 2.8	การประมาณค่าจุดภาพระหว่างจุดภาพ	13
รูปที่ 2.9	การจัดเรียงค่าสัมประสิทธิ์การแปลง.....	14
รูปที่ 2.10	วิธีการกราดค่าสัมประสิทธิ์แบบซิกแซกและแบบคู่.....	15
รูปที่ 2.11	แผนที่กลุ่มสไลซ์แบบที่นิยามด้วยพารามิเตอร์	24
รูปที่ 2.12	แผนภาพบล็อกของระบบการจำลองการสื่อสารวีดิทัศน์.....	25
รูปที่ 2.13	แผนภาพการจำลองช่องสัญญาณ	26
รูปที่ 2.14	ระยะเฉลี่ยของการเฟตติงและอัตราการตัดระดับ	27
รูปที่ 3.1	แผนภาพระบบที่นำเสนอ	35
รูปที่ 3.2	กราฟ PSNR ของอัตราการคืนสภาพมาโครบล็อกภายในที่แตกต่างกันในแต่ละอัตราบิต และที่ความถี่คอปเปิลอร์ 1 และ 10 Hz ของลำดับภาพ Sign II.....	36
รูปที่ 3.3	ขั้นตอนการทำนายการเคลื่อนที่	38
รูปที่ 3.4	แผนภาพการสร้างแผนที่การจัดสรรมาโครบล็อกแบบชัดแจ้ง	39
รูปที่ 3.5	แผนภาพการสร้างแบบจำลองบิตส่วนหัวและความซับซ้อนภายในเฟรม.....	42
รูปที่ 4.1	ค่าตัวอย่างที่ได้จากการทดลองอย่างละเอียดและการปรับเส้นโค้งของลำดับภาพ Sign I... 52	
รูปที่ 4.2	ค่าตัวอย่างที่ได้จากการทดลองอย่างละเอียดและการปรับเส้นโค้งของลำดับภาพ Sign II.. 53	

รูปที่ 4.3 ค่าตัวอย่างที่ได้จากการทดลองอย่างละเอียดและการปรับเส้นโค้งของลำดับภาพ Sign III.....	53
รูปที่ 4.4 ค่าตัวอย่างที่ได้จากการทดลองอย่างละเอียดและการปรับเส้นโค้งของลำดับภาพ Sign IV.....	54
รูปที่ 4.5 กราฟค่าเฉลี่ย PSNR เปรียบเทียบระหว่างวิธี Intra_Model กับ AAIR ของลำดับภาพ Sign IV กรณีเฟดดิ้งช้า.....	56
รูปที่ 4.6 กราฟค่าเฉลี่ย PSNR เปรียบเทียบระหว่างวิธี Intra_Model กับ AAIR ของลำดับภาพ Sign IV กรณีเฟดดิ้งเร็ว.....	56
รูปที่ 4.7 เฟรมที่ 57 ของลำดับภาพ Sign IV ที่อัตราบิต 160 K กรณีเฟดดิ้งช้าโดย (ก) Intra_Model (ข) AAIR.....	57
รูปที่ 4.8 เฟรมที่ 34 ของลำดับภาพ Sign IV ที่อัตราบิต 160 K กรณีเฟดดิ้งเร็วโดย (ก) Intra_Model (ข) AAIR.....	57
รูปที่ 4.9 กราฟค่าเฉลี่ย PSNR เปรียบเทียบระหว่างวิธี Intra_Scheme กับ Intra_Model ของลำดับภาพ Sign III กรณีเฟดดิ้งช้า.....	59
รูปที่ 4.10 กราฟค่าเฉลี่ย PSNR เปรียบเทียบระหว่างวิธี Intra_Scheme กับ Intra_Model ของลำดับภาพ Sign III กรณีเฟดดิ้งเร็ว.....	60
รูปที่ 4.11 เฟรมที่ 80 ของลำดับภาพ Sign III ที่อัตราบิต 96 K กรณีเฟดดิ้งช้าโดย (ก) Intra_Scheme (ข) Intra_Model.....	60
รูปที่ 4.12 เฟรมที่ 85 ของลำดับภาพ Sign III ที่อัตราบิต 96 K กรณีเฟดดิ้งเร็วโดย (ก) Intra_Scheme (ข) Intra_Model.....	60
รูปที่ 4.13 เฟรมที่ 67 ของลำดับภาพ Sign IV ที่อัตราบิต 160 K กรณีเฟดดิ้งช้าโดย (ก) Intra_Scheme (ข) Intra_Model.....	61
รูปที่ 4.14 เฟรมที่ 34 ของลำดับภาพ Sign IV ที่อัตราบิต 160 K กรณีเฟดดิ้งเร็วโดย (ก) Intra_Scheme (ข) Intra_Model.....	61
รูปที่ 4.15 กราฟค่าเฉลี่ย CIM เปรียบเทียบระหว่างวิธี FMO_Imp+Intra_Scheme, FMO_Dis+Intra_Scheme และ Intra_Scheme ของลำดับภาพ Sign IV กรณีเฟดดิ้งช้า.....	65

รูปที่ 4.16 กราฟค่าเฉลี่ย CIM เปรียบเทียบระหว่างวิธี FMO_Imp+Intra_Scheme, FMO_Dis+Intra_Scheme และ Intra_Scheme ของลำดับภาพ Sign I กรณีเฟตติงเร็ว.....	65
รูปที่ 4.17 เฟรมที่ 41 ของลำดับภาพ Sign IV ที่อัตราบิต 128 K กรณีเฟตติงช้าโดย (ก) FMO_Imp+Intra_Scheme (ข) FMO_Dis+Intra_Scheme (ค) Intra_Scheme	66
รูปที่ 4.18 เฟรมที่ 74 ของลำดับภาพ Sign IV ที่อัตราบิต 128 K กรณีเฟตติงเร็วโดย (ก) FMO_Imp+Intra_Scheme (ข) FMO_Dis+Intra_Scheme (ค) Intra_Scheme	66
รูปที่ 4.19 เฟรมที่ 54 ของลำดับภาพ Sign I ที่อัตราบิต 96 K กรณีเฟตติงช้าโดย (ก) FMO_Imp+Intra_Scheme (ข) FMO_Dis+Intra_Scheme (ค) Intra_Scheme	67
รูปที่ 4.20 เฟรมที่ 57 ของลำดับภาพ Sign I ที่อัตราบิต 96 K กรณีเฟตติงเร็วโดย (ก) FMO_Imp+Intra_Scheme (ข) FMO_Dis+Intra_Scheme (ค) Intra_Scheme	67
รูปที่ 4.21 เปรียบเทียบระหว่างบิตที่สร้างขึ้นจริงกับค่าจากการประมาณความซับซ้อนภายในเฟรมของลำดับภาพ Sign III ที่อัตราบิต 64K.....	70
รูปที่ 4.22 เปรียบเทียบระหว่างบิตที่สร้างขึ้นจริงกับค่าจากการประมาณความซับซ้อนภายในเฟรมของลำดับภาพ Sign II ที่อัตราบิต 160K.....	70
รูปที่ 4.23 กราฟค่าเฉลี่ย CIM เปรียบเทียบระหว่างวิธี Er_Tool+Fr_MB_RC, Er_Tool+Fr_RC และ AAIR ของลำดับภาพ Sign I กรณีเฟตติงช้า	74
รูปที่ 4.24 กราฟค่าเฉลี่ย CIM เปรียบเทียบระหว่างวิธี Er_Tool+Fr_MB_RC, Er_Tool+Fr_RC และ AAIR ของลำดับภาพ Sign I กรณีเฟตติงเร็ว	74
รูปที่ 4.25 กราฟค่าเฉลี่ย CIM เปรียบเทียบระหว่างวิธี Er_Tool+Fr_MB_RC, Er_Tool+Fr_RC และ AAIR ของลำดับภาพ Sign IV กรณีเฟตติงช้า.....	75
รูปที่ 4.26 กราฟค่าเฉลี่ย CIM เปรียบเทียบระหว่างวิธี Er_Tool+Fr_MB_RC, Er_Tool+Fr_RC และ AAIR ของลำดับภาพ Sign IV กรณีเฟตติงเร็ว.....	75
รูปที่ 4.27 เฟรมที่ 16 ของลำดับภาพ Sign II ที่อัตราบิต 64 K กรณีเฟตติงช้าโดย (ก) Er_tool+Fr_MB_RC (ข) Er_tool+Fr_RC (ค) AAIR.....	76
รูปที่ 4.28 เฟรมที่ 71 ของลำดับภาพ Sign II ที่อัตราบิต 64 K กรณีเฟตติงเร็วโดย (ก) Er_tool+Fr_MB_RC (ข) Er_tool+Fr_RC (ค) AAIR.....	76

รูปที่ 4.29 เฟรมที่ 87 ของลำดับภาพ Sign III ที่อัตราบิต 128 K กรณีสเฟดดิ้งช้าโดย (ก)
 Er_tool+Fr_MB_RC (ข) Er_tool+Fr_RC (ค) AAIR..... 77

รูปที่ 4.30 เฟรมที่ 85 ของลำดับภาพ Sign III ที่อัตราบิต 128 K กรณีสเฟดดิ้งเร็วโดย (ก)
 Er_tool+Fr_MB_RC (ข) Er_tool+Fr_RC (ค) AAIR..... 77



บทที่ 1

บทนำ

เนื้อหาในบทนี้คือภาพรวมของวิทยานิพนธ์ได้แก่ ความเป็นมาและความสำคัญของปัญหา วัตถุประสงค์ของวิทยานิพนธ์ ขอบเขตของวิทยานิพนธ์ ขั้นตอนการดำเนินการ และโครงสร้างเนื้อหา ของวิทยานิพนธ์

1.1 ความเป็นมาและความสำคัญของปัญหา

ในปัจจุบันเทคโนโลยีการสื่อสารได้มีการพัฒนาขึ้นอย่างรวดเร็ว โดยมีการนำไปใช้ร่วมกับ ข้อมูลสื่อประสม (multimedia) เช่น เสียง ภาพหรือวีดิทัศน์ ซึ่งข้อมูลที่มีขนาดใหญ่เหล่านี้จำเป็นต้อง ทำการบีบอัดสัญญาณวีดิทัศน์เพื่อให้ง่ายต่อการส่งถ่ายข้อมูล มาตรฐานการบีบอัดสัญญาณวีดิทัศน์ที่ นิยมนำมาประยุกต์ใช้กับระบบการสื่อสารในปัจจุบันได้แก่ มาตรฐาน H.264 ซึ่งเป็นมาตรฐานที่ ออกแบบให้มีประสิทธิภาพของการเข้ารหัสสูงและเหมาะสมกับการส่งผ่านโครงข่าย โปรแกรม ประยุกต์ที่นำเทคนิคการบีบอัดสัญญาณวีดิทัศน์มาใช้ในการบีบอัดข้อมูลสื่อประสมเช่น การสนทนา วีดิทัศน์ผ่านเครือข่าย (internet video chat) การแพร่สัญญาณภาพดิจิทัล (digital video broadcasting) การประชุมทางวีดิทัศน์ (Video Conferencing) วีดิทัศน์ตามคำขอ (Video on Demand) เป็นต้น เมื่อเทคโนโลยีการติดต่อสื่อสารผ่านโครงข่ายไร้สายยุค 3G เริ่มให้บริการทำให้คน หุนานวกรเข้าถึงการสื่อสารขั้นพื้นฐานได้มากขึ้น คนหุนานวกรจำนวนมากหันมาใช้โทรศัพท์วีดิทัศน์ (Video call) ในการติดต่อสื่อสารผ่านการใช้ภาษามือ (Sign language) ระหว่างกันซึ่งถือว่าเป็น โอกาสใหม่สำหรับการสื่อสารแบบเคลื่อนที่สำหรับกลุ่มคนเหล่านี้ ภาษาหลักของคนหุนานวกรคือการ สื่อสารด้วยภาษามือซึ่งใช้ลักษณะของมือทำเป็นสัญลักษณ์เพื่อใช้ในการสื่อสาร ดังนั้นโทรศัพท์วีดิทัศน์ จึงเป็นการสื่อสารที่ให้ความเป็นธรรมชาติกับคนหุนานวกรช่วยให้สามารถเห็นลักษณะท่ามือของคู่ สนทนาอีกฝั่ง นอกจากนี้การสื่อสารผ่านข้อความตัวอักษร (Text messaging) ยังถูกใช้เป็นการสื่อสาร เพิ่มเติมนอกเหนือจากการใช้ภาษามือแต่เนื่องจากการสื่อสารโดยอาศัยการพิมพ์ทำให้ช้าและไม่ สะดวกเท่ากับการสื่อสารผ่านวีดิทัศน์

ปัญหาหลักของการรับ/ส่งข้อมูลสื่อประสมผ่านโครงข่ายไร้สายคือ ช่องสัญญาณมีความไม่ คงที่ของควมกว้างช่องสัญญาณ, การสูญหายของกลุ่มข้อมูล และความผิดพลาดที่เกิดขึ้นในกลุ่ม ข้อมูล เนื่องจากลักษณะของช่องสัญญาณไร้สายมีการลดทอนของสัญญาณ ซึ่งเป็นลักษณะเฉพาะของ ช่องสัญญาณไร้สาย รายละเอียดกล่าวในหัวข้อ 2.3 ข้อมูลที่ส่งผ่านช่องสัญญาณไร้สายอาจสูญหาย หรือผิดพลาด ความผิดพลาดดังกล่าวนี้ส่งผลต่อคุณภาพของวีดิทัศน์ ดังนั้นจึงจำเป็นต้องมีวิธีการ

เข้ารหัสวีดิทัศน์ที่มีการปรับตัวได้ตามสภาวะที่เปลี่ยนแปลงของช่องสัญญาณ มีการจัดสรรบิตสำหรับวีดิทัศน์ที่เหมาะสม และสามารถแก้ไขหรือลดผลของการสูญหายหรือความผิดพลาดของกลุ่มข้อมูลที่เกิดขึ้น

มาตรฐาน H.264 มีประสิทธิภาพในการบีบอัดที่อัตราบิตต่ำแต่ยังคงคุณภาพของวีดิทัศน์ที่สูงเมื่อเปรียบเทียบกับมาตรฐานในการบีบอัดก่อนหน้านี้และมีการนิยามเครื่องมือด้านทานความผิดพลาดเพื่อลดผลกระทบของวีดิทัศน์ที่ถูกบีบอัดจากการสูญหายของข้อมูลเมื่อมีการนำไปใช้งานบนโครงข่ายไร้สาย รายละเอียดเครื่องมือด้านทานความผิดพลาดในมาตรฐานการเข้ารหัสวีดิทัศน์ H.264 กล่าวในหัวข้อ 2.2 ถึงแม้มาตรฐาน H.264 ในปัจจุบันจะประสบความสำเร็จในการบีบอัดสำหรับวีดิทัศน์ทั่วไปแต่ยังคงไม่สามารถสร้างความน่าเชื่อถือให้วีดิทัศน์ภาษามือมีความชัดเจนในการแปลความหมายเพียงพอโดยเฉพาะอย่างยิ่งที่อัตราบิตต่ำ

จึงมีหลายงานวิจัยที่นำเสนอการแก้ไขปัญหาดังกล่าว เช่น Mobile American Sign Language project (Mobile ASL) คือโครงการวิจัยระบบโทรศัพท์สำหรับคนพิการ โดยใช้การเข้ารหัสวีดิทัศน์บนโทรศัพท์เคลื่อนที่เพื่อให้คนพิการใช้ในการติดต่อสื่อสารผ่านทางภาษามือ สิ่งที่ได้นำเสนอในโครงการ MobileASL คือการหาค่าพารามิเตอร์ที่เหมาะสมเช่นอัตราเฟรมและความละเอียดเชิงพื้นที่เพื่อใช้ในการเข้ารหัสที่ทำให้วีดิทัศน์ภาษามือมีความชัดเจนในการแปลความหมายเพิ่มขึ้นสูงสุด [1-3] นอกจากนี้ยังมุ่งเน้นที่จะเข้ารหัสในบริเวณที่เกี่ยวข้องกับภาษามือตามแบบจำลองทางภาษามือ [4, 5] โดยให้ความละเอียดของวีดิทัศน์ที่สูงในบริเวณ หน้า มือและลำตัวตามลำดับและให้ความละเอียดที่ต่ำในบริเวณอื่นๆภายในเฟรมภายใต้อัตราบิตที่มีอยู่อย่างจำกัดและลดความซับซ้อนของขั้นตอนวิธีในการบีบอัดอย่างมีประสิทธิภาพเพื่อลดการใช้พลังงานของอุปกรณ์สื่อสาร

นอกจากนี้ในงานวิจัย [6] มีการใช้การคืนสภาพภายใน (intra refresh) ซึ่งเป็นหนึ่งในเครื่องมือด้านทานความผิดพลาดที่มีในมาตรฐาน H.264 เพื่อลดผลกระทบของข้อมูลที่สูญหายสำหรับประเทศไทยได้มีการพัฒนาระบบการสื่อสารภาษามือไทยอิเล็กทรอนิกส์ (eTSL-Comm) [7] วัตถุประสงค์เพื่อเป็นแพลตฟอร์มของการสื่อสารระหว่างคนหูหนวกหรือระหว่างคนหูหนวกกับคนปกติ โดยมีการประยุกต์ใช้การควบคุมอัตราในบริเวณใบหน้าและมือกับการสื่อสารผ่านวีดิทัศน์ที่อัตราบิตต่ำ

ในงานวิจัยที่เกี่ยวข้องกับการเข้ารหัสวีดิทัศน์ภาษามือก่อนหน้านี้ต่างให้ความสนใจกับการแก้ปัญหาที่เกิดจากการส่งผ่านช่องสัญญาณไร้สายโดยเฉพาะใน [6] ที่ใช้อัตราคืนสภาพภายในเพื่อยับยั้งการแพร่กระจายความผิดพลาดที่เกิดขึ้นระหว่างเฟรม เนื่องจากความผิดพลาดที่เกิดขึ้นภายในเฟรมจะนำไปสู่การแพร่กระจายอย่างไม่มีกำหนด ยิ่งข้อมูลเกิดการสูญหายในบริเวณที่เกี่ยวข้องกับ

ภาษามือโดยเฉพาะบริเวณหน้าและมือจะส่งผลกระทบต่อการใช้ใจความหมายภายในวีดิทัศน์อย่างสูงซึ่งอาจทำให้วีดิทัศน์นั้นไม่สามารถนำไปใช้ประโยชน์ได้ อย่างไรก็ตามแม้ว่าจะมีการยับยั้งการแพร่กระจายความผิดพลาดระหว่างเฟรมแต่ก็ยังมี การแพร่กระจายความผิดพลาดภายในเฟรมที่ควรถูกพิจารณา การจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้เป็นเครื่องมือด้านทานความผิดพลาดที่มีในมาตรฐาน H.264 เป็นการเข้ารหัสโดยใช้หลายกลุ่มสไลซ์ต่อเฟรมซึ่งช่วยให้ความผิดพลาดจำกัดอยู่แค่ภายในกลุ่มสไลซ์เท่านั้นไม่มีการแพร่กระจายไปยังกลุ่มสไลซ์อื่น

ในวิทยานิพนธ์นี้จึงสนใจใช้การจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้ร่วมกับการคืนสภาพภายในเพื่อช่วยให้การเข้ารหัสวีดิทัศน์สามารถด้านทานความผิดพลาดอย่างมีประสิทธิภาพ การคืนสภาพภายในทำให้เกิดการใช้บิตจำนวนมากในการเข้ารหัสเพื่อไม่ให้กระทบกับบิตโดยรวมจึงมีการใช้การคืนสภาพภายในเฉพาะบริเวณที่เกี่ยวข้องกับภาษามือตามแบบจำลองทางภาษา อย่างไรก็ตามการใช้การจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้เพื่อช่วยจำกัดการแพร่กระจายของความผิดพลาดจะส่งผลให้จำนวนบิตส่วนหัวเพิ่มขึ้น เพื่อให้ได้บิตเป้าหมายในแต่ละเฟรมตรงตามที่ตัวเข้ารหัสกำหนดจึงมีการปรับปรุงการควบคุมอัตราการระดับเฟรมโดยใช้แบบจำลองบิตส่วนหัวและการวัดความซับซ้อนภายในเฟรม นอกจากนี้เพื่อเป็นการเพิ่มความชัดเจนในการแปลความหมายในวีดิทัศน์ภาษามือเมื่อทำการเข้ารหัสที่อัตราบิตต่ำ วิทยานิพนธ์นี้จึงได้นำการควบคุมอัตราการระดับหน่วยพื้นฐานย่อยมาใช้เพื่อจัดสรรอัตราบิตให้กับบริเวณที่เกี่ยวข้องกับภาษามือตามแบบจำลองทางภาษา

1.2 วัตถุประสงค์ของวิทยานิพนธ์

พัฒนาเครื่องมือสำหรับการเข้ารหัสในมาตรฐานวีดิทัศน์ H.264 เพื่อช่วยให้ด้านทานความผิดพลาดบนช่องสัญญาณไร้สาย ด้วยการคืนสภาพมาโครบล็อกภายในร่วมกับการสร้างแผนที่กลุ่มสไลซ์เพื่อใช้สำหรับการจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้ตามแบบจำลองทางภาษา รวมถึงการปรับปรุงการควบคุมอัตราการระดับเฟรมและหน่วยพื้นฐานย่อยเพื่อเพิ่มความชัดเจนในการแปลความหมายในวีดิทัศน์ภาษามือ

1.3 ขอบเขตของวิทยานิพนธ์

1. พัฒนาเครื่องมือด้านทานความผิดพลาดโดยใช้การคืนสภาพมาโครบล็อกภายใน การจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้และการปรับปรุงการควบคุมอัตราทั้งระดับเฟรมและมาโครบล็อกในการส่งวีดิทัศน์ H.264 บนช่องสัญญาณไร้สาย
2. แบบจำลองการคืนสภาพมาโครบล็อกภายในและแผนที่กลุ่มสไลซ์แบบซัดแจ็งถูกพัฒนาโดยใช้ข้อมูลป้อนกลับจากตัวถอดรหัส

3. ใช้การจำลองการสื่อสารบนช่องสัญญาณด้วยซอฟต์แวร์จำลองช่องสัญญาณไร้สายแทนการรับ/ส่งจริงบนช่องสัญญาณไร้สาย

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1. ได้ศึกษาความรู้พื้นฐานในการเข้ารหัสสัญญาณวิทยุและการส่งสัญญาณวิทยุ
2. เข้าใจวิธีในการใช้การคืนสภาพภายใน การจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้และการควบคุมอัตราบิตสำหรับการส่งวิทยุ H.264
3. ปรับปรุงคุณภาพของวิทยุทัศน์ให้มีความเข้าใจง่ายและเหมาะสมสำหรับคนหูหนวกเพื่อใช้ในการติดต่อสื่อสารผ่านทางภาษามือภายใต้อัตราบิตที่มีอยู่อย่างจำกัด
4. สามารถนำไปประยุกต์ใช้งานกับการส่งวิทยุสัญญาณเวลาจริงหรือการประชุมทางวิทยุทัศน์ผ่านช่องสัญญาณไร้สายที่ใช้มาตรฐานการเข้ารหัสวิทยุ H.264

1.5 ขั้นตอนและวิธีการดำเนินงาน

1. ศึกษาการทำงานและขั้นตอนการเข้ารหัสของมาตรฐานการเข้ารหัสสัญญาณวิทยุ H.264
2. ศึกษาขั้นตอนและงานวิจัยที่เกี่ยวข้องกับเทคนิคการต้านทานต่อความผิดพลาดรวมกับการควบคุมอัตราของมาตรฐานการเข้ารหัสสัญญาณวิทยุ H.264
3. ศึกษาวิจัยที่เกี่ยวข้องกับคุณลักษณะเฉพาะของวิทยุทัศน์ภาษามือ
4. ออกแบบและทดสอบแบบจำลองการคืนสภาพมาโครบล็อกภายในที่ใช้ข้อมูลป้อนกลับจากตัวถอดรหัส
5. ทดสอบการใช้แบบจำลองการคืนสภาพมาโครบล็อกภายในร่วมกับการจัดเรียงมาโครบล็อกภายในแบบยืดหยุ่นได้
6. ปรับปรุงเทคนิคการควบคุมอัตราทั้งระดับเฟรมและมาโครบล็อกเมื่อมีการใช้เครื่องมือการต้านทานความผิดพลาด
7. จำลองการส่งวิทยุทัศน์ที่ถูกเข้ารหัสบนช่องสัญญาณไร้สาย
8. ทำการทดสอบ วิเคราะห์และเปรียบเทียบผลการทดสอบ

1.6 ลำดับของการนำเสนอเนื้อหา

ลำดับของการนำเสนอเนื้อหาวิทยานิพนธ์ในบทต่อไปจากบทนี้เป็นดังนี้

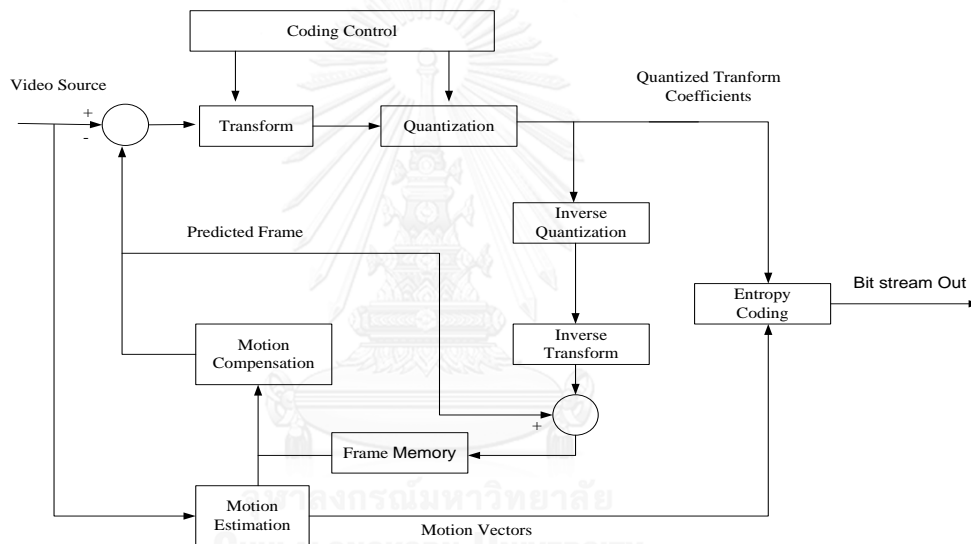
1. บทที่ 2 ความรู้พื้นฐานและทฤษฎีที่เกี่ยวข้อง เนื้อหาเป็นเรื่องความรู้พื้นฐานเกี่ยวกับการเข้ารหัสวีดิทัศน์ การควบคุมอัตราสำหรับการเข้ารหัสวีดิทัศน์ เครื่องมือด้านทานความผิดพลาดแบบจำลองช่องสัญญาณไร้สาย โครงสร้างทางภาษามือและงานวิจัยที่เกี่ยวข้อง
2. บทที่ 3 เทคนิคที่นำเสนอ เนื้อหาอธิบายการใช้เทคนิคที่ช่วยให้ด้านทานความผิดพลาดร่วมกับการควบคุมอัตราทั้งระดับเฟรมและมาโครบล็อกสำหรับการใช้งานกับมาตรฐาน H.264 ที่นำเสนอในวิทยานิพนธ์นี้
3. บทที่ 4 ผลการจำลอง เนื้อหาเป็นผลการจำลองที่ได้จากเทคนิคที่นำเสนอเมื่อทำการเข้ารหัสวีดิทัศน์ที่อัตราบิตและความผิดพลาดที่แตกต่างกันไป
4. บทที่ 5 สรุปผลการจำลองและข้อเสนอแนะ เนื้อหากล่าวถึงข้อสรุปที่ได้จากการจำลองการสื่อสารวีดิทัศน์ไร้สาย ข้อดีและข้อเสียของวิธีที่ได้นำเสนอ



บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง

2.1 มาตรฐานการเข้ารหัสสัญญาณวิดีโอ H.264 [8]

หลักการพื้นฐานของมาตรฐานการเข้ารหัสสัญญาณวิดีโอ H.264 มีความคล้ายคลึงกับมาตรฐานการเข้ารหัสสัญญาณวิดีโอทั่วไปคือ ใช้หลักการลดความซ้ำซ้อน (Redundancy) ทั้งเชิงพื้นที่ (Spatial) และเชิงเวลา (Temporal) โดยที่เข้ารหัสแบบผสม (Hybrid Coding) คือการเข้ารหัสที่มีการลดความซ้ำซ้อนทั้งในเชิงพื้นที่และในเชิงเวลา สำหรับหลักการเข้ารหัสสัญญาณวิดีโอ H.264 สามารถแสดงบล็อกไดอะแกรมตามรูปที่ 2.1



รูปที่ 2.1 บล็อกไดอะแกรมการเข้ารหัสวิดีโอ H.264

สัญญาณวิดีโอที่เข้ามาในแต่ละเฟรมจะถูกแบ่งออกเป็นบล็อกขนาด 16×16 จุดภาพที่เรียกว่ามาโครบล็อก (Macroblock) และสามารถแบ่งให้มีขนาดเล็กลงได้อีกขึ้นอยู่กับกระบวนการเข้ารหัส ในแต่ละมาโครบล็อกจะทำการเข้ารหัสภายในเฟรม (Intra frame Coding) ซึ่งเป็นการลดความซ้ำซ้อนเชิงพื้นที่ หรือ การเข้ารหัสระหว่างเฟรม (Inter Coding) ซึ่งเป็นการลดความซ้ำซ้อนเชิงเวลา ขึ้นอยู่กับการกำหนดชนิดของรูปภาพในแต่ละเฟรมที่จะทำการเข้ารหัส ถ้าเป็นเฟรมชนิด I (Intra Frame) จะทำการเข้ารหัสภายในเฟรมเท่านั้น แต่ถ้าเป็นเฟรมชนิด P (Predicted Frame) หรือเฟรมชนิด B (Bidirectional Interpolated Frame) จะทำการเข้ารหัสระหว่างเฟรม

แต่ละมาโครบล็อกในเฟรมจะใช้การแปลงจำนวนเต็ม (Integer Transform) เพื่อลดความซ้ำซ้อนเชิงพื้นที่ ซึ่งจะทำการภายในบล็อกขนาด 4×4 จุดภาพ โดยการแปลงจะแปลงจากโดเมนทางปริภูมิ

(Spatial Domain) ให้อยู่ในรูปแบบของโดเมนทางความถี่ (Frequency Domain) จากนั้นจะทำการควอนไทซ์ (Quantization) สัมประสิทธิ์ที่ได้จากการแปลงเพื่อบีบอัดข้อมูล ผลจากการควอนไทซ์ จะเกิดการสูญเสียในเชิงพื้นที่เกิดขึ้นขึ้นอยู่กับคุณภาพที่ต้องการหลังถอดรหัส การควบคุมอัตรา (Rate Control) ถูกนำมาใช้เพื่อหาค่าควอนไทซ์ที่เหมาะสมที่สุดและได้อัตราบิตตามที่ต้องการ จากนั้นค่าสัมประสิทธิ์ของการควอนไทซ์จะถูกทำให้อยู่ในรูปหนึ่งมิติด้วยวิธีการสแกนแบบซิกแซก (Zigzag Scan) และนำไปลดความซ้ำซ้อนของข้อมูลด้วยการเข้ารหัสเอนโทรปี (Entropy Coding)

สัญญาณที่ผ่านการแปลงและการควอนไทซ์แล้ว จะทำการแปลงค่าควอนไทซ์ย้อนกลับ (Inverse Quantization) และทำการแปลงผกผันจำนวนเต็ม (Inverse Integer Transform) ให้อยู่ในรูปแบบของโดเมนทางเวลาซึ่งถูกนำไปใช้เป็นเฟรมอ้างอิง (Reference Frame) สำหรับขั้นตอนการประมาณการเคลื่อนที่ (Motion Estimation) และขั้นตอนการชดเชยการเคลื่อนที่ (Motion Compensation) ซึ่งการประมาณการเคลื่อนที่จะได้เวกเตอร์การเคลื่อนที่ (Motion Vector) ที่ถูกนำไปเข้ารหัสเอนโทรปี ส่วนขั้นตอนการชดเชยการเคลื่อนที่ (Motion Compensation) ซึ่งเฟรมที่ได้จากการชดเชยการเคลื่อนที่จะนำไปใช้ในการหาเฟรมส่วนเหลือ (Residual Frame) ซึ่งจะนำไปเข้ารหัสต่อไปของการเข้ารหัสระหว่างเฟรม

2.1.2 ชนิดของรูปภาพ

ชนิดของรูปภาพแบ่งตามลักษณะของการเข้ารหัสสามารถแบ่งได้ 3 ชนิด ได้แก่ I, P และ B โดยมีรายละเอียดดังนี้

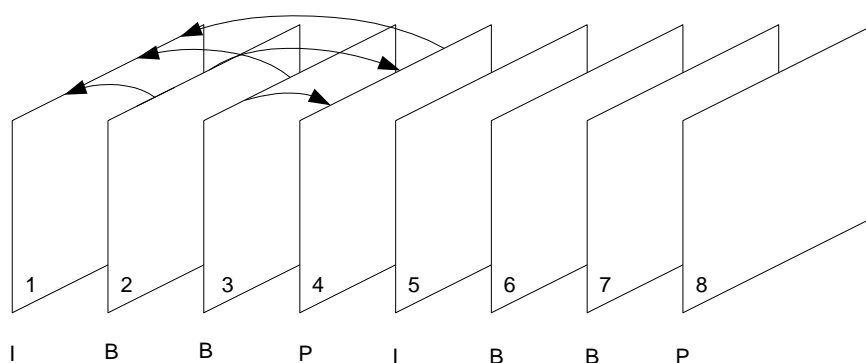
1. ภาพชนิด I (Intra) ทุกมาโครบล็อกภายในเฟรมจะถูกเข้ารหัสโดยใช้การทำนายแบบภายในเฟรม ซึ่งเป็นการเข้ารหัสโดยไม่ต้องอาศัยข้อมูลจากเฟรมอื่น ดังนั้นลำดับภาพของวิดีโอที่ค้นพบเฟรมแรกจะเป็นภาพแบบ I เสมอ โดยเป็นเฟรมที่ถูกบีบอัดภายในเฟรม ภาพแบบ I มีอัตราการบีบอัดข้อมูลต่ำที่สุดเมื่อเทียบกับการเข้ารหัสภาพทั้ง 3 แบบ

2. ภาพชนิด P (Predicted) มาโครบล็อกภายในเฟรมจะถูกเข้ารหัสโดยใช้การทำนายแบบภายในหรือระหว่างเฟรมก็ได้ โดยส่วนใหญ่จะใช้การทำนายระหว่างเฟรม เฟรมชนิด P ถูกเข้ารหัสและถอดรหัสโดยอาศัยข้อมูลของเฟรมก่อนหน้า ซึ่งเฟรมก่อนหน้าที่ใช้เรียกว่าเฟรมอ้างอิงโดยอาจจะเป็น I หรือ P เฟรมก็ได้ สำหรับการอ้างอิงถึงข้อมูลของเฟรมอ้างอิงต้องระบุเวกเตอร์การเคลื่อนที่มาด้วย ผลของการบีบอัดภาพแบบ P มีขนาดประมาณร้อยละ 30 ถึง 50 ของเฟรมแบบ I

3. ภาพชนิด B (Bidirectional interpolated) เป็นภาพที่ถูกเข้ารหัสและถอดรหัสโดยอาศัยข้อมูลทั้งจากเฟรมก่อนหน้าและเฟรมถัดไปรวมกับค่าการเคลื่อนที่ของเฟรมทั้ง 2 ดังนั้นเฟรมชนิดนี้จะไม่ถูกใช้เป็นเฟรมอ้างอิง กล่าวคือภาพชนิด B ถัดไปจะไม่ใช้ B เฟรมก่อนหน้าหรือ B เฟรม

ถัดไปในการอ้างอิง ภาพชนิดนี้ให้การบีบอัดข้อมูลสูงสุด กล่าวคือ มีขนาดเหลือเพียงร้อยละ 15 ถึง 25 ของเฟรมแบบ I เท่านั้น

โดยภาพชนิด I มีการใช้จำนวนบิตมากที่สุด รองลงมาคือภาพชนิด P และภาพชนิด B ตามลำดับ ความสัมพันธ์ระหว่างภาพชนิด I, P และ B ตามรูปที่ 2.2 ภาพเฟรมลำดับที่ 4 ซึ่งเป็นภาพแบบ P ได้จากการอ้างอิงเฟรมลำดับที่ 1 (เฟรมแบบ I) ส่วนเฟรมแบบ B ในลำดับที่ 2 และ 3 ได้จากการอ้างอิงเฟรมแบบ I และ P



รูปที่ 2.2 ความสัมพันธ์ระหว่างเฟรมแบบ I P และ B ในการทำนายการเคลื่อนที่

2.1.2 การเข้ารหัสภายในเฟรม

การเข้ารหัสชนิดนี้เป็นการลดความซ้ำซ้อนเชิงพื้นที่ (Spatial Redundancy) ของเฟรมด้วยการบีบอัดข้อมูลพื้นฐานการพิจารณาความซ้ำซ้อนของข้อมูลภายในหนึ่งเฟรมโดยเทคนิคต่างๆดังนี้

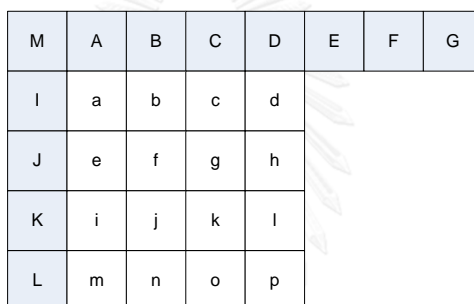
1. การสุ่มข้อมูล (Sub sampling) คือ การเข้ารหัสสัญญาณวีดิทัศน์แบบบางส่วน กล่าวคือ การเข้ารหัสแบบบางจุดภาพหรือเข้าด้วยค่าเฉลี่ยของจุดภาพเท่านั้น ทำให้สามารถลดข้อมูลลงได้แต่มีผลเสีย คือ ความคมชัดของภาพลดลง เพราะ ข้อมูลบางส่วนที่ไม่ได้ทำการเข้ารหัสต้องถูกสร้างขึ้นใหม่โดยการประมาณ(Interpolation)

2. การควอนไทซ์แบบหยาบหรือการลดความลึกของจุดภาพ (Coarse Quantization or Depth Reduction) เป็นการลดจำนวนข้อมูลโดยข้อมูลบางส่วนของภาพถูกทิ้งไป

3. การเข้ารหัสแบบการแปลง (Transform coding) ซึ่งเป็นการแปลงข้อมูลภาพจากโดเมนปริภูมิ (Spatial domain) เป็นโดเมนความถี่ (Frequency Domain) สัมประสิทธิ์จากการแปลงจะอยู่ในรูปสองมิติและถูกจัดให้อยู่มิติเดียวเพื่อทำการเข้ารหัสต่อไป

4. การทำควอนไทซ์แบบเวกเตอร์ (Vector Quantization) ข้อมูลภาพจะถูกแบ่งกลุ่มและแสดงแทนด้วยสัญลักษณ์จากชุดรหัส (code book) เพื่อทำการส่งสัญลักษณ์แทนเท่านั้น ซึ่งเป็นการประหยัดจำนวนบิตในการส่ง

สำหรับมาตรฐานการเข้ารหัสสัญญาณวิดีโอ H.264 การเข้ารหัสภายในเฟรมจะแบ่งออกเป็นบล็อกส่วนประกอบความส่องสว่างขนาด 4×4 จุดภาพหรือบล็อกส่วนประกอบความส่องสว่างขนาด 16×16 จุดภาพ เมื่อใช้การเข้ารหัสภายในเฟรมบล็อกขนาด 4×4 จุดภาพ ข้อมูลภายในบล็อกถูกทำนายจากข้อมูลใกล้เคียงเชิงพื้นที่ตามรูปที่ 2.3 ค่าตัวอย่างทั้ง 16 ค่า (จุดภาพ a ถึง p) จะถูกทำนายโดยใช้ค่าตัวอย่างที่ถูกถอดรหัสแล้วก่อนหน้าของบล็อกที่อยู่ติดกัน สำหรับบล็อกขนาด 4×4 จุดภาพ รองรับโหมดการทำนายทั้งสิ้น 9 โหมดใน 8 ทิศทาง โหมดการทำนายทั้ง 9 โหมด ตามรูปที่ 2.4



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

รูปที่ 2.3 ตำแหน่งส่วนประกอบความส่องสว่างขนาด 4×4 จุดภาพ

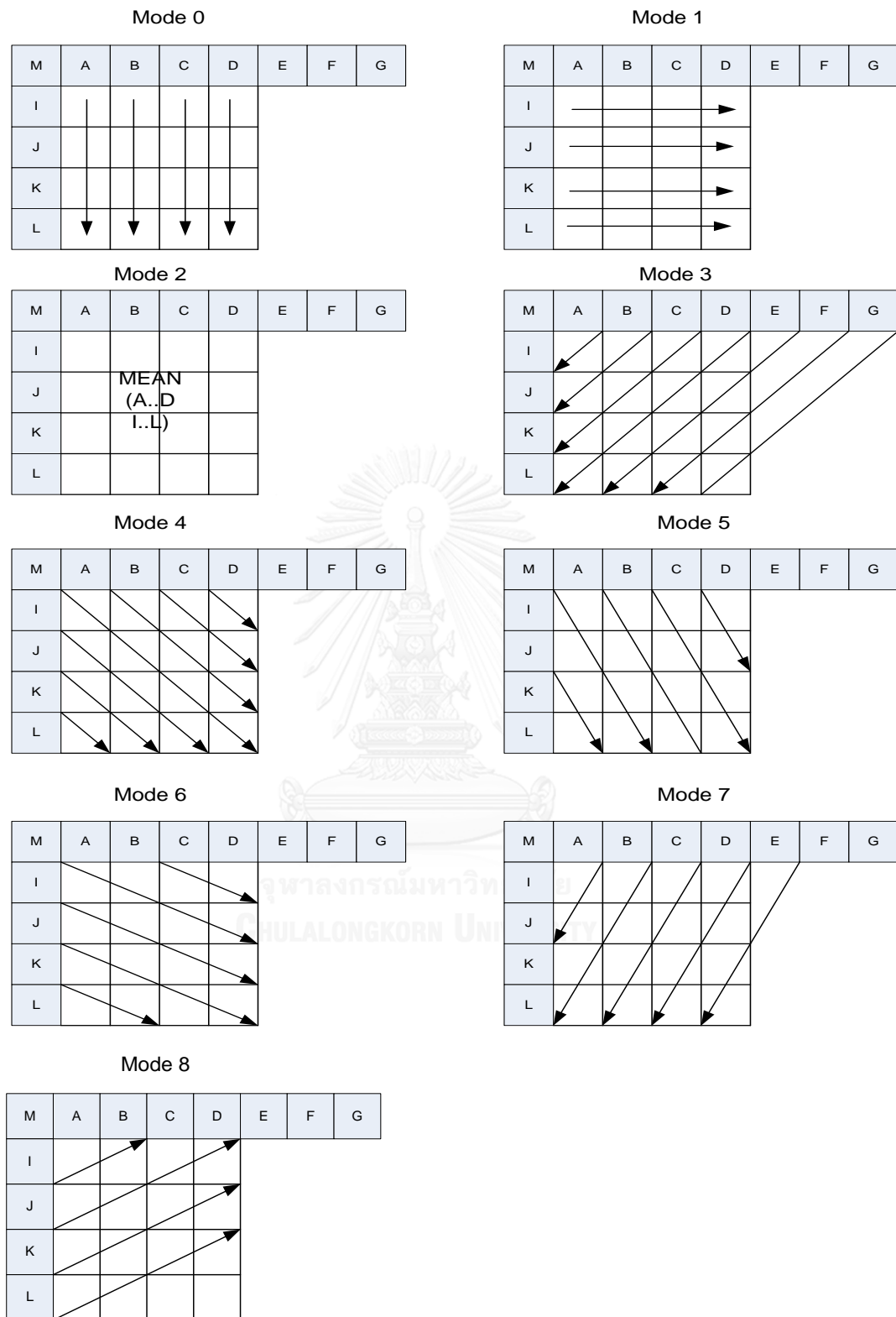
โดยแต่ละโหมดมีรายละเอียดดังนี้

โหมด 0 เป็นการทำนายในแนวตั้ง ค่าตัวอย่างจากด้านบนของบล็อกขนาด 4×4 จุดภาพจะถูกคัดลอกลงมาบล็อกด้านล่างตามที่ระบุโดยลูกศร

โหมด 1 เป็นการการทำนายในแนวนอน มีการทำนายคล้ายกับโหมด 0 แตกต่างกันตรงที่จะคัดลอกค่าตัวอย่างจากบล็อกด้านซ้ายของบล็อกขนาด 4×4 จุดภาพ

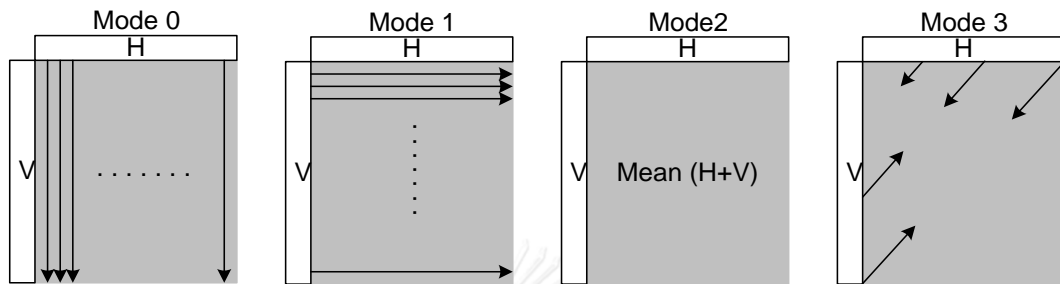
โหมด 2 คือโหมดการทำนายแบบดีซี (DC prediction) เป็นการทำนายโดยเฉลี่ยจากค่าตัวอย่างที่อยู่ติดกัน

โหมด 3 – 8 เป็นโหมดการทำนายแนวทแยงมุมจาก ซ้าย-ล่าง, ขวา-ล่าง, แนวตั้ง-ขวา, แนวนอน-ล่าง, แนวตั้ง-ซ้ายและแนวนอน-บน ตามลำดับ



รูปที่ 2.4 รูปแบบการเข้ารหัสภายในเฟรมของส่วนประกอบความส่องสว่างขนาด 4x4 จุดภาพ

สำหรับบล็อกส่วนประกอบความส่องสว่างขนาด 16×16 จุดภาพ มีทั้งหมด 4 โหมดการทำนาย โดยโหมด 0 - 2 คือ โหมดการทำนายในแนวตั้ง, โหมดการทำนายในแนวนอน, โหมดการทำนายแบบ DC โดยจะมีการทำนายคล้ายกับบล็อกขนาด 4×4 จุดภาพ แต่จะแตกต่างกันตรงที่เปลี่ยนเป็นใช้บล็อกขนาด 16×16 จุดภาพ ส่วนโหมดที่ 3 ทำนายจากค่าตัวอย่างในแนวนอนลงมาซ้าย และจากค่าตัวอย่างในแนวตั้งขึ้นไปขวา ตามรูปที่ 2.5



รูปที่ 2.5 รูปแบบการเข้ารหัสภายในเฟรมของส่วนประกอบความส่องสว่างขนาด 16×16 จุดภาพ

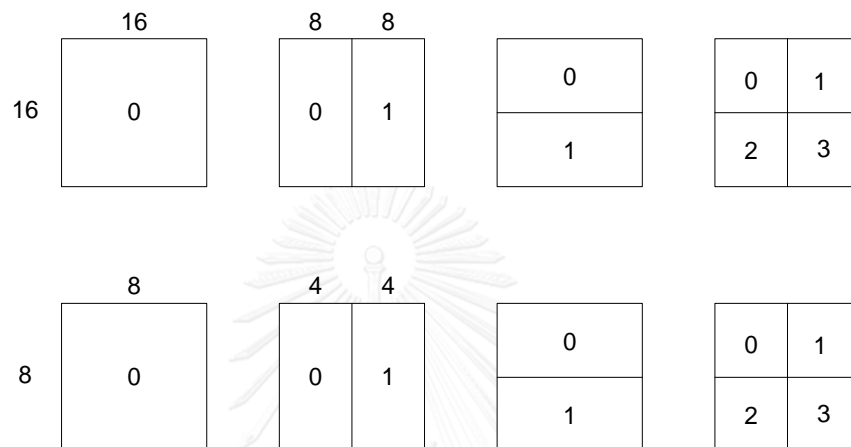
สำหรับบล็อกส่วนประกอบความเข้มสีขนาด 8×8 จุดภาพ จะถูกทำนายโดยใช้เทคนิคคล้ายคลึงกับบล็อกส่วนประกอบความส่องสว่างขนาด 16×16 เนื่องจากบล็อกส่วนประกอบความส่องสว่างมักจะมี ความราบเรียบ (Smooth) ทั่วบริเวณกว้าง

โดยการเลือกโหมดของการเข้ารหัสในแต่ละแบบจะพิจารณาจากการหาค่าความแตกต่างสัมบูรณ์รวม (Sum Absolute Difference) ที่มีค่าน้อยที่สุด

2.1.3 การเข้ารหัสระหว่างเฟรม

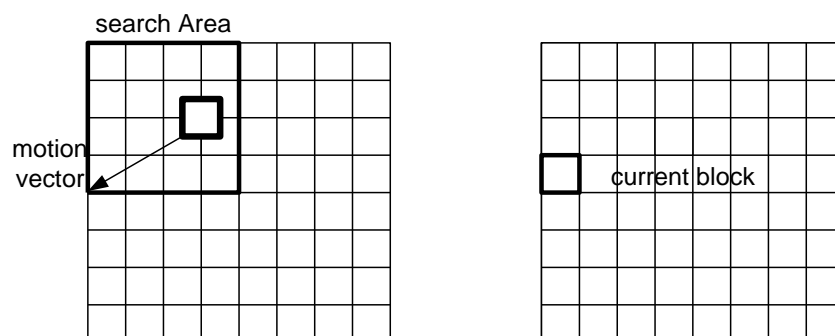
สำหรับสัญญาณวิดีโอที่เฟรมที่อยู่ติดกันจะมีความสัมพันธ์ระหว่างข้อมูลเชิงเวลา ซึ่งทำให้เกิดความซ้ำซ้อนของข้อมูลจึงทำให้สามารถเข้ารหัสเฉพาะความแตกต่างได้ ในการลดข้อมูลที่มีความซ้ำซ้อนเหล่านี้สามารถใช้เทคนิคการประมาณการเคลื่อนที่และการชดเชยการเคลื่อนที่เพื่อลดข้อมูลการประมาณการเคลื่อนที่เป็นการหาความแตกต่างระหว่างจุดภาพที่สอดคล้องกันระหว่างเฟรมสองเฟรมที่ติดกัน โดยทั่วไปเราสามารถหาค่าตัวอย่างของจุดภาพในภาพที่มีการเคลื่อนที่ของวัตถุได้เลื่อนจากจุดหนึ่งในเฟรมหนึ่งไปยังจุดอื่นในเฟรมถัดไป ถ้าเราสามารถหาตำแหน่งที่เปลี่ยนแปลงหรือที่เรียกว่า เวกเตอร์การเคลื่อนที่ซึ่งการประมาณการเคลื่อนที่แม่นยำสามารถช่วยทำให้บีบอัดข้อมูลได้มากขึ้น วิธีการประมาณการเคลื่อนที่แบบย้อนกลับเป็นการสร้างข้อมูลจุดภาพด้วยเวกเตอร์การเคลื่อนที่และข้อมูลจุดภาพอื่นร่วมกับความแตกต่างของจุดภาพเรียกว่า การชดเชยการเคลื่อนที่

โดยมาโครบล็อกชนิด P เฟรมจะถูกแบ่งสัดส่วนเพื่อใช้ในการชดเชยการเคลื่อนที่ มาตรฐานการเข้ารหัส H.264 สามารถแบ่งบล็อกส่วนประกอบความส่องสว่างขนาด 16×16 จุดภาพออกเป็นบล็อกขนาดต่างๆได้ 4 แบบ คือ 16×16 , 16×8 , 8×16 และ 8×8 จุดภาพ ในกรณีที่บล็อกขนาด 8×8 จุดภาพถูกเลือก บล็อกขนาด 8×8 จุดภาพสามารถแบ่งออกเป็นบล็อกขนาดต่างๆได้อีก 4 แบบคือ 8×8 , 8×4 , 4×8 และ 4×4 จุดภาพ เหมือนกันกับบล็อกส่วนประกอบความเข้มสี การแบ่งบล็อกขนาด 16×16 จุดภาพและบล็อกขนาด 8×8 จุดภาพ ตามรูปที่ 2.6



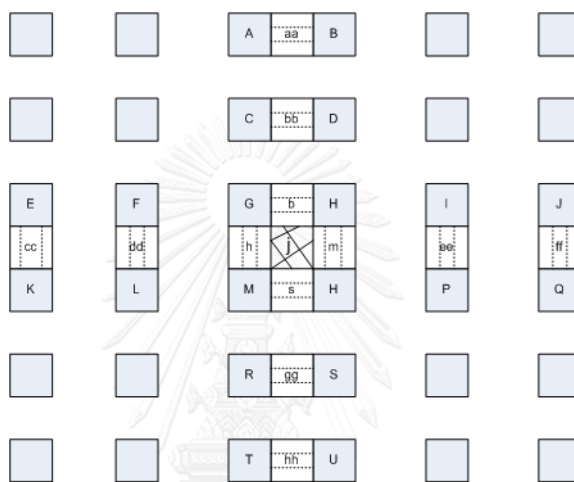
รูปที่ 2.6 แถวบน : การแบ่งบล็อกขนาด 16×16 จุดภาพ, แถวล่าง : การแบ่งบล็อกขนาด 8×8 จุดภาพ

การหาเวกเตอร์การเคลื่อนที่ที่ทำโดยหาบล็อกที่มีจุดภาพใกล้เคียงมากที่สุดระหว่างเฟรมปัจจุบันและเฟรมก่อนหน้าหรือเรียกว่าเฟรมอ้างอิง การค้นหาบล็อกที่ใกล้เคียงจะกำหนดบริเวณการค้นหา ซึ่งอาจเป็นทั้งเฟรมหรือเป็นเพียงบางส่วนของเฟรม วิธีที่ใช้ในการวัดความใกล้เคียงของบล็อกทั้ง 2 มีหลายวิธี เช่น Mean Absolute Difference (MAD), Sum Square Difference (SSD) และ Sum Absolute Difference (SAD) ตามรูปที่ 2.7 พร้อมเวกเตอร์การเคลื่อนที่



รูปที่ 2.7 การประมาณการเคลื่อนที่

เพื่อความแม่นยำสูงในการประมาณการเคลื่อนที่ มาตรฐานการเข้ารหัส H.264 ใช้การประมาณการเคลื่อนที่ละเอียดในระดับ 1/2 และ 1/4 จุดภาพ การหาค่าระหว่างจุดภาพที่ 1/2 จุดภาพ สามารถคำนวณได้จากการใช้วงจรกรองตอบสนองอิมพัลส์จำกัด (Finite Impulse Response, FIR) 6 แถบในแนวตั้งและแนวนอน ในขณะที่การหาค่าระหว่างจุดภาพที่ 1/4 จุดภาพ สามารถคำนวณได้จากการเฉลี่ยค่าจุดภาพที่จำนวนเต็มและ 1/2 จุดภาพ การสร้างค่าจุดภาพสร้างจากการชดเชยการเคลื่อนที่เมื่อใช้การประมาณการเคลื่อนที่ในความละเอียดน้อยกว่าระดับจุดภาพแสดงตามรูปที่ 2.8

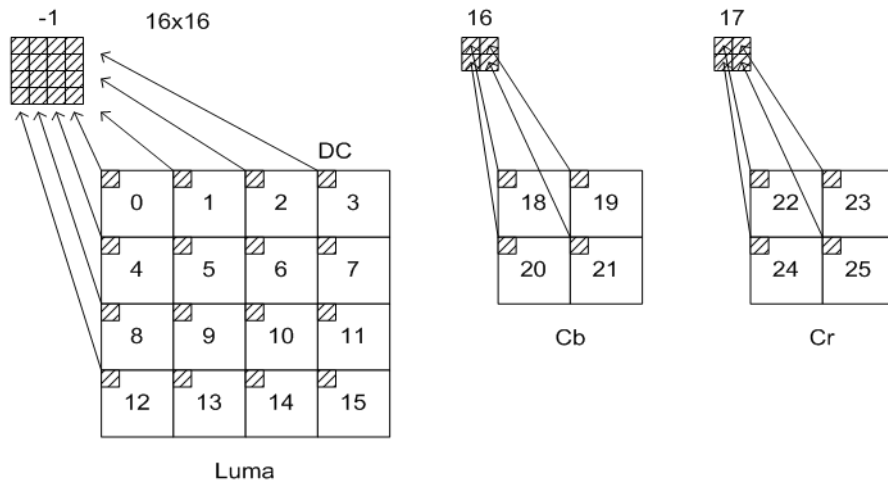


รูปที่ 2.8 การประมาณหาค่าจุดภาพระหว่างจุดภาพ

นอกจากโหมดสำหรับเข้ารหัสมาโครบล็อกชนิด P เฟรมที่กล่าวมาข้างต้นแล้ว มาโครบล็อกสามารถถูกเข้ารหัสโดยใช้โหมดข้าม (Skip mode) ได้ โดยการใช้โหมดการเข้ารหัสนี้ ค่าความผิดพลาดจากการทำนาย, เวกเตอร์การเคลื่อนที่และพารามิเตอร์ดัชนีอ้างอิงจะไม่ถูกส่ง

2.1.4 การแปลงและการแปลงผกผัน

มาตรฐาน H.264 ใช้การแปลงแบบจำนวนเต็มขนาดบล็อก 4x4 จุดภาพ ซึ่งมีพื้นฐานจากการแปลงโคไซน์วิญุต (Discrete Cosine Transform, DCT) ซึ่งกรณีนี้ทำการแปลงผกผันของการแปลงจำนวนเต็มจะมีประสิทธิภาพดีกว่าไม่มีความผิดพลาดจากการแปลงที่ไม่เป็นจำนวนเต็ม หมายเลขตำแหน่งสำหรับการเรียงลำดับของสัมประสิทธิ์การแปลงในบล็อกตามรูปที่ 2.9



รูปที่ 2.9 การจัดเรียงค่าสัมประสิทธิ์การแปลง

การแปลงสำหรับกลุ่มจุดภาพรับเข้า X ขนาด 4×4 จุดภาพคำนวณได้ตามสมการที่ (2.1) ดังนี้

$$Y = AXA^T \quad (2.1)$$

$$A = \begin{bmatrix} a & a & a & a \\ b & c & -c & -b \\ a & -a & -a & a \\ c & -b & b & -c \end{bmatrix}$$

เมื่อ

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

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

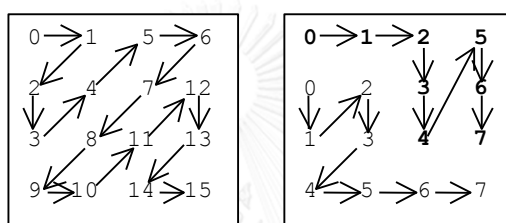
$$c = \frac{1}{\sqrt{2}} \cos\left(\frac{3\pi}{8}\right)$$

2.1.5 การแจกหน่วยและการกราดค่าสัมประสิทธิ์

การแจกหน่วยเป็นขั้นตอนที่เกิดการสูญเสียความละเอียดข้อมูลบางส่วน ทำให้การเข้ารหัสเกิดเป็นการเข้ารหัสคงข้อมูลหลักแทนที่จะเป็นการเข้ารหัสคงรายละเอียด การสูญเสียความละเอียดข้อมูลขึ้นอยู่กับค่าแจกหน่วยคือ ถ้าค่าแจกหน่วยมีค่ามากการสูญเสียข้อมูลมาก ค่าแจกหน่วยมีค่าน้อยการสูญเสียข้อมูลน้อย การเลือกค่าแจกหน่วยเป็นขั้นตอนที่มีความสำคัญต่อจำนวนบิตที่ทำการ

เข้ารหัสเฟรมซึ่งเป็นกระบวนการที่อยู่ในขั้นตอนของการควบคุมอัตรา มาตรฐาน H.264 กำหนดระดับชั้นบันไดของการแจกหน่วยไว้ที่ 52 ระดับ [16]

เมื่อผ่านการกราดค่าสัมประสิทธิ์ของการแปลงแบบจำนวนเต็ม เรียกค่าสัมประสิทธิ์แรก ด้านบนซ้ายมือซึ่งเป็นของส่วนประกอบความถี่ต่ำสุด (ปกติมีค่ามากที่สุด) ว่าค่าสัมประสิทธิ์ DC ส่วนสัมประสิทธิ์อื่นที่เหลือนอกจากสัมประสิทธิ์ DC เรียกว่าสัมประสิทธิ์ AC เมื่อผ่านการแจกหน่วยใช้การกราดเป็นการเรียงค่าสัมประสิทธิ์ให้เป็นมิติเดียว โดยเริ่มจากค่าสัมประสิทธิ์ DC ซึ่งในมาตรฐาน H.264 นี้จะมีวิธีเรียงค่าสัมประสิทธิ์อยู่ 2 วิธีคือ วิธีแรกคือการกราดแบบซิกแซก เป็นวิธีที่ใช้อยู่โดยทั่วไปทั้งในมาตรฐานก่อนหน้า และอีกวิธีคือ การกราดแบบคู่ ลักษณะการเรียงค่าสัมประสิทธิ์ทั้ง 2 วิธีตามรูปที่ 2.10



รูปที่ 2.10 วิธีกราดค่าสัมประสิทธิ์แบบซิกแซกและแบบคู่

2.1.6 การเข้ารหัสเอนโทรปี

เป็นขั้นตอนการแปลงค่าจากสัมประสิทธิ์ทางความถี่ที่ได้จากการแจกหน่วยให้เป็นเลขฐานสองเพื่อส่งเป็นกระแสบิตและเป็นอีกขั้นตอนหนึ่งในการบีบอัดข้อมูลเพื่อลดความซ้ำซ้อนที่อยู่ในกระแสบิต มาตรฐาน H.264 มีวิธีการเข้ารหัสแบบเอนโทรปี 2 วิธี คือ Context-Adaptive Variable Length Coding (CAVLC) และ Context-Based Adaptive Arithmetic Coding (CABAC) โดยวิธี CABAC สามารถลดอัตราบิตได้มากกว่า CAVLC ประมาณร้อยละ 5 ถึง 15

2.1.7 การควบคุมอัตราสำหรับการเข้ารหัสสัญญาณวิดีโอที่ต้นชั้นสูง

การควบคุมอัตราสำหรับ AVC แบ่งออกได้ 3 ระดับ คือ ระดับกลุ่มภาพ (GOP level) ระดับภาพ (Picture level) และระดับพื้นฐานย่อย (Basic unit level) โดยที่ในระดับกลุ่มภาพทำการคำนวณจำนวนบิตสำหรับเข้ารหัสกลุ่มภาพนั้นๆ จากนั้นในระดับภาพ ทำการคำนวณหาจำนวนบิตเป้าหมายที่ทำการเข้ารหัสในแต่ละภาพ ส่วนในระดับหน่วยพื้นฐานย่อย จะทำการหาจำนวนบิตเป้าหมายในหน่วยพื้นฐานย่อยนั้นๆ เมื่อได้จำนวนบิตเป้าหมายในระดับหน่วยพื้นฐานย่อยแล้วจะทำการหาค่าควอนไทซ์พารามิเตอร์โดยใช้โมเดลกำลังสอง ซึ่งการหาค่าควอนไทซ์พารามิเตอร์นี้ใช้ค่าความแตกต่างสัมบูรณ์ (Mean Absolute Difference : MAD) ด้วย จากนั้นทำการตรวจสอบเงื่อนไข

เพื่อกำหนดค่าควอนไทซ์พารามิเตอร์ เมื่อเข้ารหัสครบทุกหน่วยพื้นฐานย่อยแล้วทำการรวมจำนวนบิตทั้งหมดภายในภาพ แล้วทำการคำนวณหาจำนวนบิตสะสมคงเหลือในบัฟเฟอร์

2.1.7.1 การควบคุมอัตราการระดับกลุ่มภาพ (GOP level rate control)

การควบคุมอัตราการระดับกลุ่มภาพคำนวณบิตทั้งหมดสำหรับภาพที่เหลืออยู่ใน GOP และค่าควอนไทซ์พารามิเตอร์เริ่มต้นของภาพแรกที่ถูกเก็บสะสม (stored) เมื่อภาพที่ j^{th} ในกลุ่มภาพ i^{th} ถูกเข้ารหัส บิตทั้งหมดสำหรับภาพที่เหลืออยู่ใน GOP สามารถคำนวณได้จากสมการที่ (2.2)

$$B_i(j) = \begin{cases} \frac{R_i(j)}{f} \times N_i - V_i(j) & j=1 \\ B_i(j-1) + \frac{R_i(j) - R_i(j-1)}{f} \times (N_i - j + 1) - b_i(j-1) & j=2,3,\dots,N_i \end{cases} \quad (2.2)$$

1) สำหรับภาพแรกใน GOP ($j=1$) บิตทั้งหมดสามารถคำนวณได้จากสมการที่ (2.2) โดยที่ f คืออัตราเฟรม

N_i คือจำนวนของภาพทั้งหมดใน i^{th} GOP

$R_i(j)$ คืออัตราบิตที่สามารถใช้ได้

$V_i(j)$ คือบิตสะสมในบัฟเฟอร์

2) สำหรับภาพอื่นๆ บิตทั้งหมดสามารถคำนวณได้จากสมการที่ (2.2) ในกรณีที่อัตราบิตคงที่ การคำนวณบิตทั้งหมดสามารถคำนวณได้จากสมการที่ (2.3)

$$B_i(j) = B_i(j-1) - b_i(j-1) \quad (2.3)$$

โดยที่ $b_i(j-1)$ คือบิตที่สร้างขึ้นจริงใน $(j-1)^{th}$

3) $V_i(j)$ จะถูกปรับหลังจากเข้ารหัสแต่ละภาพตามสมการที่ (2.4)

$$V_i(1) = \begin{cases} 0 & i=1 \\ V_{i-1}(N_{i-1}) & \text{other} \end{cases} \quad (2.4)$$

$$V_i(j) = V_i(j-1) + b_i(j-1) - \frac{R_i(j-1)}{f} \quad j=2,3,\dots,N_i$$

2.1.7.2 การควบคุมอัตราการระดับเฟรม (Frame level rate control)

ในระดับนี้แบ่งการทำงานได้ เป็น 2 ขั้นตอนคือ

1) **Pre-Encoding Stage** ทำการคำนวณหาจำนวนบิตเป้าหมายเพื่อเข้ารหัสในแต่ละเฟรม แต่สำหรับกรณีที่เข้ารหัสเพียง 2 ระดับ คือระดับกลุ่มภาพและระดับเฟรม ในขั้นตอนนี้จะคำนวณหา

ค่าควอนไทล์พารามิเตอร์ด้วย โดยวิธีที่แตกต่างกันจะถูกนำไปใช้กับเฟรมที่ถูกเก็บสะสมและไม่ถูกเก็บสะสม

1.1) เฟรมที่ไม่ถูกเก็บสะสม (Non-stored picture)

ค่าควอนไทล์พารามิเตอร์จะถูกคำนวณโดยใช้ทฤษฎีการประมาณค่าในช่วง (Interpolation) ดังต่อไปนี้ สมมติว่า j^{th} และ $(j+L+1)^{th}$ เฟรมถูกเก็บสะสมและค่าควอนไทล์พารามิเตอร์สำหรับเฟรมที่ถูกเก็บสะสมนั้นคือ $QP_i(j)$ และ $QP_i(j+L+1)$ ตามลำดับ ค่าควอนไทล์พารามิเตอร์ของ i^{th} เฟรมที่ไม่ถูกเก็บสะสมสามารถถูกคำนวณตาม 2 กรณีดังต่อไปนี้

กรณีที่ 1 เมื่อ $L = 1$ มีเฉพาะ 1 เฟรมที่ไม่ถูกเก็บสะสมระหว่าง 2 เฟรมที่ถูกเก็บสะสม ค่าควอนไทล์พารามิเตอร์สามารถคำนวณได้จากสมการที่ (2.5)

$$QP_i(j+1) = \begin{cases} \frac{QP_i(j) + QP_i(j+2) + 2}{2} & \text{if } QP_i(j) \neq QP_i(j+2) \\ QP_i(j) + 2 & \text{Otherwise} \end{cases} \quad (2.5)$$

กรณีที่ 2 เมื่อ $L > 1$ มีมากกว่า 1 เฟรมที่ไม่ถูกเก็บสะสมระหว่าง 2 เฟรมที่ถูกเก็บสะสม ค่าควอนไทล์พารามิเตอร์สามารถคำนวณได้จากสมการที่ (2.6)

$$QP_i(j+k) = QP_i(j) + \alpha + \max\left\{\min\left\{\frac{QP_i(j+L+1) - QP_i(j)}{L-1}, 2 \times (k-1)\right\}, -2 \times (k-1)\right\} \quad (2.6)$$

โดยที่ $k = 1, \dots, L$ และ α ถูกกำหนดโดย

$$\alpha = \begin{cases} -3 & QP_i(j+L+1) - QP_i(j) \leq -2 \times L - 3 \\ -2 & QP_i(j+L+1) - QP_i(j) = -2 \times L - 2 \\ -1 & QP_i(j+L+1) - QP_i(j) = -2 \times L - 1 \\ 0 & QP_i(j+L+1) - QP_i(j) = -2 \times L \\ 1 & QP_i(j+L+1) - QP_i(j) = -2 \times L + 1 \\ 2 & \text{Otherwise} \end{cases} \quad (2.7)$$

1.2) เฟรมที่ถูกเก็บสะสม (Stored picture)

ค่าควอนไทล์พารามิเตอร์ของเฟรมที่ถูกเก็บสะสมถูกคำนวณผ่าน 2 ขั้นตอนดังต่อไปนี้

ขั้นตอนที่ 1 กำหนดบิตเป้าหมายสำหรับแต่ละเฟรมที่ถูกเก็บสะสม

- 1) กำหนดบิตเป้าหมายสำหรับเฟรมที่ถูกเก็บสะสมใน GOP ปัจจุบัน

ระดับบัพเฟอร์เป้าหมายถูกกำหนดล่วงหน้าสำหรับแต่ละเฟรมที่ถูกจัดเก็บตามบิตที่ใช้ในการเข้ารหัสของเฟรมแรกใน GOP 1^{th} กับเฟรมแรกที่ถูกจัดเก็บและค่าเฉลี่ยของความซับซ้อนในเฟรม

หลังจากเข้ารหัสเฟรมแรกที่ถูกเก็บสะสมใน i^{th} GOP ค่าเริ่มต้นของระดับบัพเฟอร์เป้าหมายถูกกำหนดให้เป็นตามสมการที่ (2.8)

$$S_i(2) = V_i(2) \quad (2.8)$$

ระดับบัพเฟอร์เป้าหมายสำหรับเฟรมที่ถูกเก็บสะสมต่อมาสามารถคำนวณได้จากสมการที่ (2.9)

$$S_i(j+1) = S_i(j) - \frac{S_i(2)}{N_p(i)-1} + \frac{\bar{W}_{p,i}(j) \times (L+1) \times R_i(j)}{f \times (\bar{W}_{p,i}(j) + \bar{W}_{b,i}(j) \times L)} - \frac{R_i(j)}{f} \quad (2.9)$$

โดยที่ $\bar{W}_{p,i}(j)$ คือค่าเฉลี่ยถ่วงน้ำหนักความซับซ้อนของเฟรมที่ถูกเก็บสะสม

$\bar{W}_{b,i}(j)$ คือค่าเฉลี่ยถ่วงน้ำหนักความซับซ้อนของเฟรมที่ไม่ถูกเก็บสะสม

ค่าเฉลี่ยถ่วงน้ำหนักเหล่านี้สามารถคำนวณได้จากสมการที่ (2.10)

$$\begin{aligned} \bar{W}_{p,i}(j) &= \frac{W_{p,i}(j)}{8} + \frac{7 \times \bar{W}_{p,i}(j-1)}{8} \\ \bar{W}_{b,i}(j) &= \frac{W_{b,i}(j)}{8} + \frac{7 \times \bar{W}_{b,i}(j-1)}{8} \\ W_{p,i}(j) &= b_i(j) \times QP_{p,i}(j) \\ W_{b,i}(j) &= \frac{b_i(j) \times QP_{b,i}(j)}{1.3636} \end{aligned} \quad (2.10)$$

เมื่อไม่มีเฟรมที่ถูกเก็บสะสมระหว่าง 2 เฟรมที่ถูกเก็บสะสมจากสมการที่ (2.9) สามารถคำนวณได้จากสมการที่ (2.11)

$$S_i(j+1) = S_i(j) - \frac{S_i(2)}{N_p(i)-1} \quad (2.11)$$

2) คำนวณบิตเป้าหมายสำหรับเฟรมที่ถูกเก็บสะสมปัจจุบัน

บิตเป้าหมายถูกจัดสรรสำหรับ j^{th} เฟรมที่ถูกเก็บสะสมใน i^{th} GOP ถูกกำหนดโดยขึ้นอยู่กับระดับบัพเฟอร์เป้าหมาย (2.9) อัตราเฟรม ความกว้างช่องสัญญาณและบิตสะสมในบัพเฟอร์สามารถคำนวณได้จากสมการที่ (2.12)

$$\tilde{T}_i(j) = \frac{R_i(j)}{f} + \gamma \times (S_i(j) - V_i(j)) \quad (2.12)$$

โดยที่ γ คือค่าคงที่มีค่าเท่ากับ 0.5 เมื่อไม่มีเฟรมที่ถูกเก็บสะสมและมีค่าเท่ากับ 0.25 ในกรณีอื่นๆ

ในขณะเดียวกันจำนวนของบิตที่เหลืออยู่ควรถูกพิจารณาเมื่อบิตเป้าหมายถูกคำนวณจากสมการที่ (2.13)

$$\hat{T}_i(j) = \frac{W_{p,i}(j-1) \times B_i(j)}{W_{p,i}(j-1) \times N_{p,r} + W_{b,i}(j-1) \times N_{b,r}} \quad (2.13)$$

โดยที่ $N_{p,r}$ และ $N_{b,r}$ คือจำนวนของเฟรมที่ถูกเก็บสะสมที่ยังเหลืออยู่และจำนวนของเฟรมที่ไม่ถูกเก็บสะสมที่ยังเหลืออยู่ตามลำดับ

บิตเป้าหมายคือการรวมกันของ $\tilde{T}_i(j)$ และ $\hat{T}_i(j)$ ที่ถูกถ่วงน้ำหนักสามารถคำนวณได้จากสมการที่ (2.14)

$$T_i(j) = \beta \times \hat{T}_i(j) + (1 - \beta) \times \tilde{T}_i(j) \quad (2.14)$$

โดยที่ β คือค่าคงที่และโดยทั่วไปมีค่าเท่ากับ 0.5 เมื่อไม่มีเฟรมที่ถูกเก็บสะสมและมีค่าเท่ากับ 0.25 ในกรณีอื่นๆ

ขั้นตอนที่ 2 คำนวณค่าควอนไทซ์พารามิเตอร์และนำไปใช้กับ RDO

ค่า MAD ของเฟรมที่ถูกเก็บสะสม $\tilde{\sigma}_i(j)$ ถูกทำนายโดยใช้แบบจำลองเชิงเส้นในสมการที่ (2.15) โดยใช้ค่า MAD ของเฟรมที่ถูกเก็บสะสมก่อนหน้า $\sigma_i(j-1-L)$

$$\tilde{\sigma}_i(j) = a_1 \times \sigma_i(j-1-L) + a_2 \quad (2.15)$$

โดยที่ ค่าเริ่มต้นของค่าสัมประสิทธิ์ a_1 และ a_2 เท่ากับ 1 และ 0 ตามลำดับ โดยจะถูกปรับค่าโดยใช้สมการการถดถอยเชิงเส้นหลังจากแต่ละเฟรมหรือหน่วยพื้นฐานย่อยถูกเข้ารหัส

ค่าควอนไทซ์จะสอดคล้องกับบิตเป้าหมายโดยคำนวณจากโมเดลกำลังสองสามารถคำนวณได้จากสมการที่ (2.16)

$$T_i(j) = c_1 \times \frac{\tilde{\sigma}_i(j)}{Q_{step,i}(j)} + c_2 \times \frac{\tilde{\sigma}_i(j)}{Q_{step,i}^2(j)} - m_{h,i}(j) \quad (2.16)$$

โดยที่ $m_{h,i}(j)$ คือจำนวนของบิตส่วนหัว ในขณะที่ c_1 และ c_2 คือค่าสัมประสิทธิ์

เพื่อให้คุณภาพสัญญาณวีดิทัศน์ราบเรียบต่อเนื่องกันมากขึ้นจึงทำการปรับค่าควอนไทซ์พารามิเตอร์ดังแสดงในสมการที่ (2.17)

$$QP_i(j) = \min\{QP_i(j-L-1)+2, \max\{QP_i(j-L-1)-2, QP_i(j)\}\} \quad (2.17)$$

ค่าควอนไทซ์พารามิเตอร์จะถูกนำไปใช้กับ RDO สำหรับแต่ละมาโครบล็อกในเฟรมปัจจุบัน

2) ส่วน Post-Encoding Stage ทำการปรับค่าพารามิเตอร์ต่างๆ เช่น ในโมเดลเชิงเส้น โมเดลกำลังสอง พารามิเตอร์ที่ใช้สำหรับคำนวณค่าควอนไทล์

2.1.7.3 การควบคุมอัตราการระดับหน่วยพื้นฐานย่อย (Basic unit level rate control)

ในหนึ่งเฟรมจะประกอบไปด้วย N_{mbpic} มาโครบล็อก โดยกำหนดให้หน่วยพื้นฐานย่อยเป็น กลุ่มของมาโครบล็อกที่อยู่ติดกันจะประกอบไปด้วย N_{mbunit} มาโครบล็อก ดังนั้นจำนวนหน่วยพื้นฐานย่อยในหนึ่งเฟรม N_{unit} สามารถคำนวณได้จากสมการที่ (2.18)

$$N_{unit} = \frac{N_{mbpic}}{N_{mbunit}} \quad (2.18)$$

การควบคุมอัตราการระดับหน่วยพื้นฐานย่อยคือหาค่าควอนไทล์ที่เหมาะสมของทุกหน่วยพื้นฐานย่อยภายในเฟรม โดยที่ผลรวมของจำนวนบิตที่ใช้ในแต่ละเฟรมต้องใกล้เคียงกับจำนวนบิตเป้าหมาย $T_i(j)$ ในแต่ละเฟรมมากที่สุด โดยมี 5 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 ทำนายค่า MAD ($\tilde{\sigma}_{l,i}(j)$) ของหน่วยพื้นฐานย่อยที่เหลืออยู่ในเฟรมที่ถูกเก็บสะสม โดยใช้ MAD ที่ได้จริงของหน่วยพื้นฐานย่อยในเฟรมที่ถูกเก็บสะสมก่อนหน้า

ขั้นตอนที่ 2 คำนวณจำนวนของบิต สำหรับ I^{th} ในขั้นตอนนี้ประกอบด้วย 3 ขั้นตอนย่อยดังนี้

ขั้นตอนที่ 2.1 คำนวณบิตเป้าหมายสำหรับ I^{th} หน่วยพื้นฐานย่อย

ให้ T_r คือจำนวนของบิตที่เหลืออยู่สำหรับเฟรมปัจจุบันและค่าเริ่มต้นถูกกำหนดให้เป็น $T_i(j)$ บิตเป้าหมายสำหรับ I^{th} หน่วยพื้นฐานย่อยสามารถคำนวณได้จากสมการที่ (2.19)

$$\tilde{b}_l = T_r \times \frac{\tilde{\sigma}_{l,i}^2(l)}{\sum_{k=1}^{N_{unit}} \tilde{\sigma}_{k,i}^2(j)} \quad (2.19)$$

ขั้นตอนที่ 2.2 คำนวณค่าเฉลี่ยของจำนวนบิตส่วนหัวที่ถูกสร้างจากการเข้ารหัสหน่วยพื้นฐานย่อยได้จากสมการที่ (2.20)

$$\begin{aligned} \tilde{m}_{hdr,l} &= \tilde{m}_{hdr,l-1} \times \left(1 - \frac{1}{l}\right) + \frac{\hat{m}_{hdr,l}}{l} \\ m_{hdr,l} &= \tilde{m}_{hdr,l} \times \frac{l}{N_{unit}} + m_{hdr,l} \times \left(1 - \frac{l}{N_{unit}}\right); 1 \leq l \leq N_{unit} \end{aligned} \quad (2.20)$$

โดยที่ $\hat{m}_{hdr,l}$ คือจำนวนของบิตส่วนหัวที่ถูกสร้างขึ้นโดย I^{th} หน่วยพื้นฐานย่อยในเฟรมที่ถูกเก็บสะสม

$m_{hdr,l}$ คือการประมาณจำนวนของบิต สำหรับ I^{th} หน่วยพื้นฐานย่อย

ขั้นตอนที่ 2.3 คำนวณจำนวนของบิต \hat{b}_l สำหรับ l^{th} หน่วยพื้นฐานย่อยได้จากสมการที่ (2.21)

$$\hat{b}_l = \tilde{b}_l - m_{ndr,l} \quad (2.21)$$

ขั้นตอนที่3 คำนวณค่าควอนไทล์สำหรับ l^{th} หน่วยพื้นฐานย่อยของ j^{th} เฟรมใน i^{th} GOP โดยใช้แบบจำลองสมการกำลังสอง โดยพิจารณาได้ 3 กรณี

กรณีที่1 หน่วยพื้นฐานย่อยแรกในเฟรมปัจจุบันสามารถคำนวณได้จากสมการที่ (2.22)

$$QP_{l,i}(j) = Q\bar{P}_i(j-L-1) \quad (2.22)$$

โดยที่ คือค่าเฉลี่ยของค่าควอนไทล์พารามิเตอร์สำหรับทุกหน่วยพื้นฐานย่อยในเฟรมที่ถูกเก็บสะสม

กรณีที่2 เมื่อจำนวนของบิตที่เหลืออยู่น้อยกว่า 0 ค่าควอนไทล์พารามิเตอร์ควรมากกว่าของหน่วยพื้นฐานย่อยก่อนหน้า ดังนั้นบิตที่ถูกสร้างขึ้นจะใกล้เคียงกับบิตเป้าหมาย สามารถคำนวณได้จากสมการที่ (2.23)

$$QP_{l,i}(j) = QP_{l-1,i}(j) + \Delta_{Bu} \quad (2.23)$$

โดยที่ Δ_{Bu} คือช่วงที่เปลี่ยนแปลงของค่าควอนไทล์พารามิเตอร์ตามหน่วยพื้นฐานย่อย ค่าเริ่มต้นกำหนดให้เท่ากับ 1

กรณีที่3 ในกรณีอื่น จะคำนวณหาค่าควอนไทล์โดยใช้แบบจำลองสมการกำลังสองตามสมการ แล้วแปลงเป็นค่าควอนไทล์พารามิเตอร์ ถูกกำหนดขอบเขตตามสมการที่ (2.24)

$$QP_{l,i}(j) = \max\{QP_{l-1,i}(j) - \Delta_{Bu}, \min\{QP_{l,i}(j), QP_{l-1,i}(j) + \Delta_{Bu}\}\} \quad (2.24)$$

ขั้นตอนที่4 นำไปใช้กับ RDO สำหรับทุกมาโครบล็อกในหน่วยพื้นฐานย่อย

ขั้นตอนที่5 ปรับจำนวนของบิตที่เหลืออยู่ สัมประสิทธิ์ของแบบจำลองการทำนายเชิงเส้นและแบบจำลองสมการกำลังสอง

2.1.8 อัตราบิด – ความเพี้ยนที่เหมาะสม (Rate-distortion Optimization)

วัตถุประสงค์ของอัตราบิด-ความเพี้ยนที่เหมาะสมคือเพื่อลดค่าความเพี้ยนให้ได้น้อยที่สุดภายใต้อัตราที่มีอย่างจำกัด สถานะปัญหาของอัตราบิดความเพี้ยนที่เหมาะสมตามสมการที่ (2.25)

$$\min_{\phi} D(\phi) \text{ Subj. to } R(\phi) \leq R_c \quad (2.25)$$

โดยที่ ϕ คือชุดค่าของการตัดสินใจ $D(\phi)$ และ $R(\phi)$ คือค่าความเพี้ยนและอัตราที่เหมาะสมโดยใช้ ϕ ตามลำดับ และ R_c คือข้อจำกัดของอัตรา ปัญหาอัตราที่มีอย่างจำกัดโดยทั่วไปจะถูกแก้ปัญหโดยใช้วิธีตัวคูณลากรานจ์ (Lagrange multipliers method) โดยแก้สถานะของปัญหาใน (2.15) ให้อยู่ในรูปตามสมการที่ (2.26)

$$\min_{\phi} J(\phi; \lambda) = D(\phi) + \lambda R(\phi) \quad (2.26)$$

โดยที่ $J(\phi; \lambda)$ คือค่าใช้จ่ายลากรานจ์ (Lagrange cost) และ λ คือตัวคูณลากรานจ์ (Lagrange multiplier) โดยทั่วไปสามารถแก้ปัญหสมการที่ โดยเลือกค่า λ และ ϕ ที่ให้ค่าใช้จ่ายลากรานจ์ ($J(\phi; \lambda)$) น้อยที่สุดซึ่งการเลือกค่าตัวแปร λ และ ϕ จะทำให้ได้ค่าความเพี้ยนน้อยที่สุดภายในอัตราที่จำกัด

อัตราบิด - ความเพี้ยนที่เหมาะสมถูกนำมาประยุกต์ใช้กับการเข้ารหัสวิดีโอในระดับหน่วยพื้นฐานย่อย โดยการเลือกค่าพารามิเตอร์ในการเข้ารหัสสำหรับแต่ละมาโครบล็อกที่ให้ค่าความเพี้ยนน้อยที่สุดภายในอัตราที่มีอยู่อย่างจำกัด ค่าพารามิเตอร์ต่างๆในการเข้ารหัสได้แก่ พารามิเตอร์การควอนไทซ์ (QP) และโหมดการเข้ารหัส (p) ค่าพารามิเตอร์ที่เหมาะสมจะถูกเลือกเพื่อให้ค่าใช้จ่ายลากรานจ์มีค่าน้อยที่สุด ตามสมการที่ (2.27)

$$\min_{p, QP} J(X, p, QP; \lambda) = D(X, p, QP) + \lambda R(X, p, QP) \quad (2.27)$$

โดยที่ λ คือตัวคูณลากรานจ์

X คือตำแหน่งของมาโครบล็อก

R คืออัตราบิดที่ใช้สำหรับการเข้ารหัสมาโครบล็อกเมื่อใช้ p และ QP

D คือ ค่าความเพี้ยนที่ได้จากเข้ารหัส

โดยทั่วไปค่า λ จะสามารถหาได้ผ่านความสัมพันธ์ของค่าพารามิเตอร์การควอนไทซ์หรือการควบคุมอัตรา

2.2 เทคนิคการต้านทานความผิดพลาดในมาตรฐานการเข้ารหัส H.264/AVC

เพื่อลดผลกระทบของความผิดพลาดและการแพร่กระจายความผิดพลาดที่เกิดขึ้นจากการส่งผ่านช่องสัญญาณ มาตรฐาน H.264 ได้มีการพัฒนาเครื่องมือต้านทานความผิดพลาดขึ้นมาเพื่อใช้งานซึ่งในบางส่วนของเครื่องมือต้านทานความผิดพลาดได้มีการใช้งานในมาตรฐานก่อนหน้านี้ [8, 9]

2.2.1 ลำดับกลุ่มสไลซ์แบบกำหนดเอง (Arbitrary Slice Order: ASO)

มาตรฐาน H.264/AVC ออกแบบให้กลุ่มสไลซ์ในแต่ละเฟรมสามารถถูกส่ง รับและถอดรหัสได้อย่างอิสระโดยไม่ขึ้นกับสไลซ์อื่นภายในเฟรม เมื่อแต่ละกลุ่มสไลซ์ภายในเฟรมถูกส่งผ่านในเส้นทางที่แตกต่างกันตามสถานการณ์ที่เหมาะสมของช่องสัญญาณ ถ้าที่ฝั่งรับตัวถอดรหัสสามารถรับกลุ่มสไลซ์ล่างได้ก่อนกลุ่มสไลซ์บน ตัวถอดรหัสสามารถถอดรหัสกลุ่มสไลซ์ล่างได้เลยโดยไม่จำเป็นต้องรอกลุ่มสไลซ์บน

2.2.2 กลุ่มสไลซ์ซ้ำซ้อน (Redundant Slices)

เป็นเครื่องมือต้านทานความผิดพลาดที่อนุญาตให้ตัวเข้ารหัสใช้กลุ่มสไลซ์ซ้ำซ้อนในการเข้ารหัสโดยจะเป็นเพียงแค่บางส่วนหรือทั้งเฟรมก็ได้ นอกจากนี้กลุ่มสไลซ์ซ้ำซ้อนสามารถเข้ารหัสโดยใช้เฟรมอ้างอิงที่แตกต่างจากเฟรมดั้งเดิมในขั้นตอนการทำนายได้ ประโยชน์ของวิธีนี้ก็คือนี่คือเมื่อกลุ่มสไลซ์หลักภายในเฟรมเกิดการสูญหายตัวถอดรหัสสามารถใช้ข้อมูลจากกลุ่มสไลซ์ซ้ำซ้อนในการถอดรหัสได้ หรือถ้าเฟรมอ้างอิงที่ใช้ในขั้นตอนการทำนายที่ตัวถอดรหัสสูญหายกลุ่มสไลซ์หลักสามารถใช้ข้อมูลจากกลุ่มสไลซ์ซ้ำซ้อนซึ่งทำนายจากเฟรมอ้างอิงที่สูญหายไปในการเข้ารหัสทดแทนได้

2.2.3 การตัดแบ่งข้อมูล (Data Partitioning)

ข้อมูลของกลุ่มสไลซ์สามารถที่จะแยกองค์ประกอบวากยสัมพันธ์ (syntax) ของส่วนที่สำคัญมาก เช่น ข้อมูลส่วนหัวหรือเวกเตอร์การเคลื่อนที่ให้อยู่ในกลุ่มข้อมูลที่แตกต่างกับส่วนที่สำคัญน้อยกว่า เช่น ส่วนเหลือจากการเข้ารหัสทั้งแบบภายในและภายนอก ดังนั้นเมื่อกลุ่มข้อมูลที่ไม่สำคัญเกิดการสูญหายการถอดรหัสยังสามารถทำได้เนื่องจากข้อมูลที่สำคัญไม่ได้สูญหายไปด้วย

2.2.4 การคืนสภาพภายใน (Intra Refreshing)

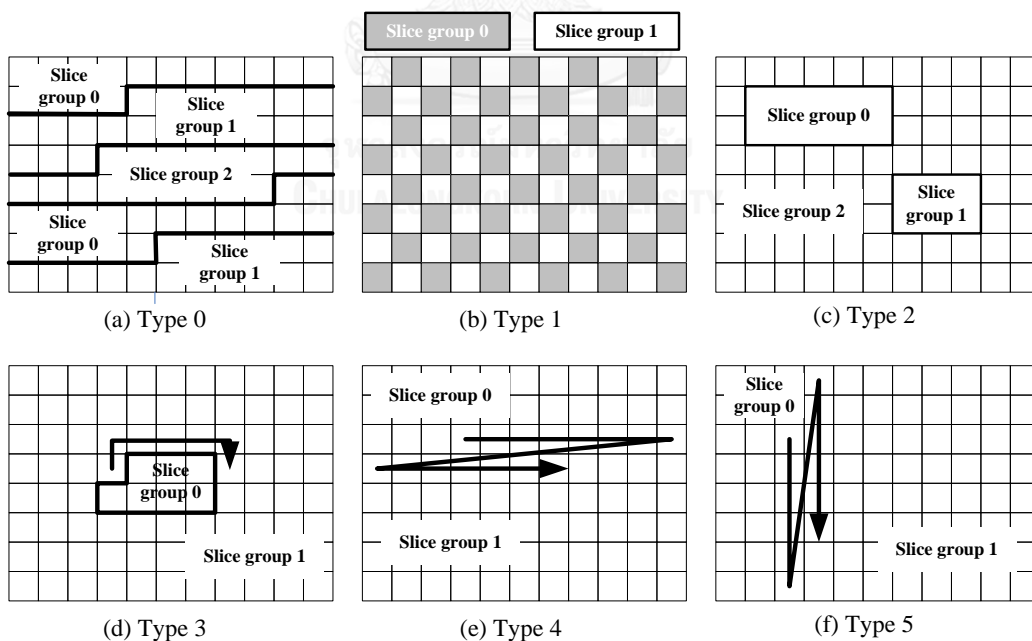
การคืนสภาพภายในเป็นเทคนิคพื้นฐานในการหลีกเลี่ยงการแพร่กระจายของความผิดพลาดเชิงเวลาเนื่องจากใช้ข้อมูลเชิงพื้นที่ของบล็อกข้างเคียงในกระบวนการทำนายในการเข้ารหัส ซึ่งทำให้มีอัตราการบีบอัดข้อมูลต่ำสุด การใช้การคืนสภาพภายในทั้งเฟรมมีการใช้บิตเป็นจำนวนมากซึ่งเป็นสาเหตุของอัตราบิตที่การแปรผันอย่างรุนแรง มาตรฐาน H.264 จึงอนุญาตให้มีการใช้การคืนสภาพ

ภายในเฉพาะบางกลุ่มสไลซ์หรือบางมาโครบล็อกภายในเฟรมซึ่งให้ประสิทธิภาพที่สูงกว่าใช้การคืนสภาพภายในทั้งเฟรม

2.2.5 การจัดเรียงมาโครบล็อกแบบยืดหยุ่น (Flexible Macroblock Ordering: FMO)

การจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้เป็นส่วนหนึ่งของขั้นตอนวิธีที่ใช้ในวิทยานิพนธ์นี้ โดยถูกพัฒนาขึ้นมาใหม่ซึ่งไม่ได้ออกแบบมาให้อยู่ในกระบวนการเข้ารหัสตามมาตรฐาน H.264 แต่ใช้เป็นเครื่องมือตัวเลือก การจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้เป็นหนึ่งในเครื่องมือที่ช่วยให้วิดีโอที่ถูกรีบอัดด้านทานต่อความผิดพลาดซึ่งการจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้ช่วยในการจัดลำดับของการเข้ารหัสมาโครบล็อกให้มีความยืดหยุ่น จากเดิมที่มีลำดับการกวาดแบบแรสเตอร์ (Raster scan) โดยใช้แผนที่การจัดสรรมาโครบล็อก (Macroblock allocation map : MBAmP) ในการจัดกลุ่มของมาโครบล็อกเข้าเป็นกลุ่มสไลซ์ (Slice group) ในแผนที่จะระบุหมายเลขของแต่ละมาโครบล็อกว่าจะถูกจัดสรรอยู่ในกลุ่มสไลซ์ใดซึ่งแต่ละกลุ่มสไลซ์จะมีการถอดรหัสที่เป็นอิสระต่อกัน

จำนวนของกลุ่มสไลซ์ถูกจำกัดให้มีจำนวนไม่เกิน 8 กลุ่มสไลซ์ต่อ 1 เฟรมเพื่อหลีกเลี่ยงแผนกการจัดสรรที่ซับซ้อน มาตรฐานการเข้ารหัส H.264 รองรับ FMO ทั้งหมด 7 ประเภท โดยประเภทที่ 0 ถึง 5 ถูกกำหนดไว้ตามมาตรฐาน ตามรูปที่ 2.11 ส่วนในประเภทที่ 6 ถูกเรียกว่า “Explicit” คือจัดให้แต่ละมาโครบล็อกอยู่ในกลุ่มสไลซ์ใดก็ได้ด้วยการกำหนดเองจากผู้ใช้ที่ตัวเข้ารหัส



รูปที่ 2.11 แผนที่กลุ่มสไลซ์แบบที่นิยามด้วยพารามิเตอร์

ถ้ามาโครบล็อกในกลุ่มสไลซ์ไม่สามารถที่จะถอดรหัสได้อันอาจเนื่องมาจากมาโครบล็อกเกิดสูญหาย ด้วยเทคนิคนี้ความผิดพลาดจะถูกแก้ไขได้ง่ายโดยใช้ความซับซ้อนเชิงพื้นที่ (Spatial

redundancy) มาโครบล็อกที่สูญหายนี้สามารถที่จะกู้คืนได้โดยใช้การประมาณค่าในช่วงเชิงพื้นที่ (Spatial interpolation) จากมาโครบล็อกที่อยู่โดยรอบในกรณีที่มาโครบล็อกโดยรอบนั้นเป็นมาโครบล็อกที่อยู่ในกลุ่มสไลซ์อื่น

การใช้การจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้ทำให้มีการแบ่งพื้นที่ออกเป็นกลุ่มสไลซ์เกิดขึ้นภายในเฟรม นอกจากจะทำให้ตัวเข้ารหัสไม่สามารถใช้การทำนายภายในเฟรมข้ามระหว่างกลุ่มสไลซ์ได้แล้วแต่ยังจำกัดเวกเตอร์การเคลื่อนที่สำหรับการถอดรหัสให้ประมาณการเคลื่อนที่ได้เฉพาะภายในกลุ่มสไลซ์เดียวกันเท่านั้น ดังนั้นจึงเป็นการลดประสิทธิภาพการเข้ารหัสเนื่องจากการจำกัดพื้นที่ในการประมาณการเคลื่อนที่ ถ้าแผนที่การจัดสรรมาโครบล็อกมีการเปลี่ยนแปลงทุกๆเฟรม บิตส่วนหัวจะถูกสร้างขึ้นและแทรกลงไปใ้ในกระแสบิต จำนวนของบิตส่วนหัวที่เพิ่มขึ้นจะเป็นการลดจำนวนบิตที่จัดสรรสำหรับการเข้ารหัสข้อมูลวิดีโอที่ส่งภายในเฟรม นอกจากนี้เป็นการลดประสิทธิภาพของการเข้ารหัสเอนโทรปีเนื่องจากการจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้จะเปลี่ยนอักขรรหัสในกรณีของ CABAC และ CAVLC ของแต่ละสไลซ์

2.3 ตัวจำลองช่องสัญญาณไร้สาย

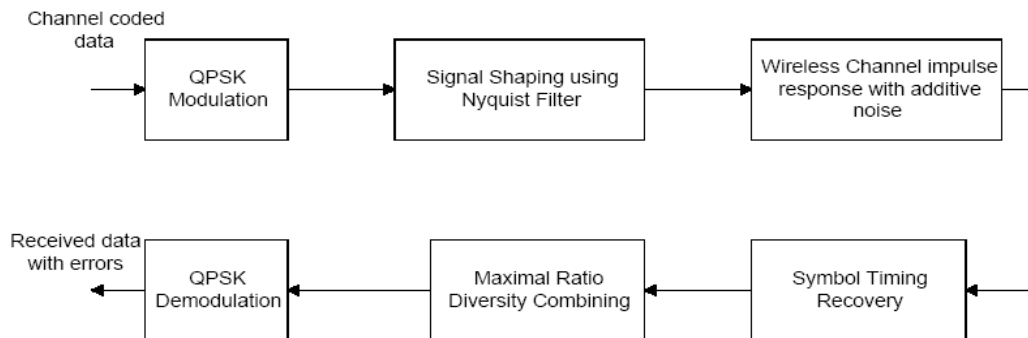
วิทยานิพนธ์นี้ทำการจำลองการสื่อสารวิดีโอที่ส่งบนช่องสัญญาณไร้สาย ตามรูปที่ 2.12



รูปที่ 2.12 แผนภาพบล็อกของระบบการจำลองการสื่อสารวิดีโอ

การวัดช่องสัญญาณไร้สายภายใต้สภาพแวดล้อมที่แตกต่างกันออกไปทำได้ยากและต้องใช้เวลาอย่างมาก อย่างไรก็ตามช่องสัญญาณไร้สายสามารถจำลองได้โดยใช้ตัวจำลองช่องสัญญาณไร้สายที่เรียกว่า “Rayleigh fading channel” [10]

คุณลักษณะของช่องสัญญาณภายใต้สภาพแวดล้อมที่แตกต่างกันสามารถศึกษาได้ผ่านการคำนวณการจำลองช่องสัญญาณไร้สายโดยการปรับเปลี่ยนค่าพารามิเตอร์ที่เหมาะสม ตัวอย่างของตัวจำลองช่องสัญญาณไร้สายที่ใช้ในวิทยานิพนธ์นี้มีรายละเอียดอยู่ใน [11] ขั้นตอนการทำงานของตัวจำลองช่องสัญญาณได้ตามรูปที่ 2.13



รูปที่ 2.13 แผนภาพการจำลองช่องสัญญาณ

2.3.1 ลักษณะเฉพาะของช่องสัญญาณไร้สาย

สัญญาณวิทยุที่ส่งผ่านช่องสัญญาณไร้สาย อาจเกิดการลดทอนของสัญญาณซึ่งจะทำให้ข้อมูลที่ส่งไปสูญหายหรือเกิดความผิดพลาดได้ ลักษณะการลดทอนของสัญญาณโดยทั่วไปสามารถแบ่งออกเป็น 3 ประเภท ได้แก่

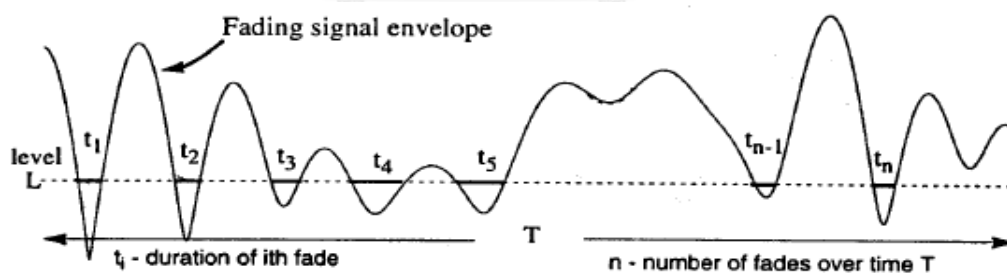
1. การสูญเสียเชิงวิถี เป็นการลดทอนของสัญญาณที่แปรตามระยะทางระหว่างภาคส่งและภาครับสัญญาณ
2. ชาโดว์อิงหรือเฟดดิ้งสเกลใหญ่ เกิดจากการที่สัญญาณบางส่วนถูกบดบังไปชั่วขณะจากสิ่งแวดล้อมรอบข้าง เช่น อาคารและเนินเขา ระดับสัญญาณที่รับได้จะเปลี่ยนแปลงในช่วงระยะเท่าๆกันของสิ่งกีดขวาง
3. เฟดดิ้งพหุวิถีหรือเฟดดิ้งสเกลเล็ก เกิดจากการที่สายอากาศของเครื่องรับสัญญาณอยู่ใกล้กับส่วนที่เป็นพื้นผิวมาก สิ่งกีดขวางต่างๆรอบข้างจึงส่งผลให้เกิดการสะท้อนและการหักเหของสัญญาณในหลายทิศทาง ส่งผลทำให้สัญญาณที่ได้รับประกอบด้วยหลายสัญญาณที่เดินทางมาจากทิศทางที่แตกต่างกันโดยมีทั้งขนาดและเฟสที่แตกต่างกัน ผลกระทบเฟดดิ้งพหุวิถีทำให้ระดับสัญญาณที่รับได้มีการเปลี่ยนแปลงบ่อยครั้ง

ประเภทของสัญญาณที่มีการเฟดดิ้งแบ่งได้โดยการคำนวณจากความสัมพันธ์ของค่าการแผ่เวลาประวิง และความกว้างช่องสัญญาณของช่องสัญญาณเป็น 2 ประเภทคือ เฟดดิ้งแบนราบและเฟดดิ้งแบบเลือกความถี่ ค่าการแผ่เวลาประวิงคือ การประวิงเวลาระหว่างสัญญาณที่มาจากเส้นทางที่ใกล้ตัวรับสัญญาณมากที่สุดกับสัญญาณที่มาจากเส้นทางที่อยู่ไกลจากตัวรับสัญญาณมากที่สุดค่าการแผ่เวลาประวิงของช่องสัญญาณ มีลักษณะเป็นตัวแปรสุ่ม จึงใช้ค่าเบี่ยงเบนมาตรฐานที่เรียกว่ารากของกำลังสองเฉลี่ย (Root Mean Square, RMS) เป็นตัววัด

ถ้าผลคูณของค่า RMS กับความกว้างช่องสัญญาณมีค่าน้อยกว่า 1 จะเรียกช่องสัญญาณนี้ว่าเป็นช่องสัญญาณที่มีเฟดดิ้งแบบแบนราบ ซึ่งในกรณีนี้กรอบคลื่นของสัญญาณที่ได้รับจะมีแอมพลิจูดและเฟสเป็นค่าสุ่ม แต่ถ้าผลคูณของค่า RMS กับความกว้างช่องสัญญาณมีค่ามากกว่า 1 จะเรียกว่าช่องสัญญาณที่มีเฟดดิ้งแบบเลือกความถี่ เฟดดิ้งแบบเลือกความถี่จะทำให้เกิดการแทรกแซงกันเองของสัญญาณ เนื่องจากค่าประวิงเวลาการแพร่มากพอที่จะทำให้สัญญาณที่มาจากหลายเส้นทางได้รับข้อมูลซ้ำซ้อนกันเอง เรียกอีกอย่างได้ว่า Intersymbol Interference, ISI

การเคลื่อนที่ของฝั่งรับและฝั่งส่งในระบบสื่อสารไร้สายจะมีผลทำให้ความถี่ของสัญญาณที่ส่งมีการเปลี่ยนแปลง ความถี่ที่เปลี่ยนแปลงไปจะเรียกว่าความถี่ดอปเปลอร์ (Doppler Frequency, f_D) ซึ่งหาได้จากอัตราส่วนระหว่างความเร็วที่เคลื่อนที่ของผู้ส่งกับความถี่ของสัญญาณที่ถูกส่งออกมา โดยความถี่ดอปเปลอร์จะมีความสัมพันธ์กับค่าเฉลี่ยของช่วงเวลาเฟดดิ้ง (Average Duration of Fade, ADF) อัตราการตัดระดับ (Level Crossing Rate : lcr) และ อัตราบิตผิดพลาด (Bit Error Rate, BER)

ค่าเฉลี่ยของ BER เป็นอัตราส่วนของผลคูณระหว่าง ADF กับ lcr ที่ระดับของพลังงานที่ตั้งไว้ ค่า ADF จะเป็นอัตราส่วนของส่วนกลับค่า f_D โดยที่ค่า lcr จะเป็นอัตราส่วนโดยตรงกับค่า f_D ค่า ADF จะสอดคล้องกับความยาวเบรสต์ของข้อมูลที่ผิดพลาด และค่า lcr จะสอดคล้องกับความถี่ที่เกิดเบรสต์ของข้อมูลที่ผิดพลาดตามรูปที่ 2.14 สัญญาณจะมีการเฟดดิ้งเมื่อกรอบคลื่นของสัญญาณอยู่ต่ำกว่าระดับ L ค่า ADF หาได้จาก $(\sum t_i) / n$ และ lcr หาได้จาก n / T



รูปที่ 2.14 ระยะเวลาเฉลี่ยของการเฟดดิ้งและอัตราการตัดระดับ

เมื่อไม่มีไดเวอร์ซิตีของสายอากาศหรือมีสายอากาศรับสัญญาณเพียงจุดเดียว อัตราของการเกิดเฟดดิ้งส่วนใหญ่จะขึ้นกับความถี่ดอปเปลอร์ การเพิ่มจำนวนสายอากาศเป็นอีกวิธีหนึ่งที่สามารถลดการเฟดดิ้งของสัญญาณได้ โดยสัญญาณที่ได้รับในแต่ละสายอากาศจะไม่มีกรรบกวนกันของสัญญาณ ซึ่งทำให้มีโอกาสที่จะได้รับสัญญาณที่มีผลกระทบของเฟดดิ้งลดลง

ในวิทยานิพนธ์นี้พิจารณาเฉพาะกรณีของเฟตติงสเกลเล็กซึ่งมีอยู่ 2 ชนิดคือช่องสัญญาณที่มีเฟตติงช้าและเฟตติงเร็วโดยขึ้นอยู่กับค่าความถี่ตอบเพลอร์ เมื่อค่าความถี่ตอบเพลอร์มีค่าน้อยจะเรียกว่าเฟตติงช้า ระยะเวลาที่เกิดเฟตติงจะยาวนานกว่าความถี่ที่เกิดเฟตติง จึงทำให้การเบิรสต์ของข้อมูลผิดพลาดมีระยะยาวสำหรับเฟตติงช้า สำหรับเฟตติงเร็วการต่อเนื่องกันของข้อมูลผิดพลาดจะมีระยะสั้นกว่าแต่จะเกิดขึ้นบ่อยครั้งกว่า เพื่อแสดงให้เห็นความชัดเจนของผลกระทบที่เกิดจึงกำหนดการจำลองช่องสัญญาณเฟตติงช้าด้วยความถี่ตอบเพลอร์ 1 Hz และช่องสัญญาณเฟตติงเร็วด้วยความถี่ตอบเพลอร์ 40 Hz รายละเอียดและผลการจำลองช่องสัญญาณอยู่ในบทที่ 4 หัวข้อ 4.1.1

2.3.2 รูปแบบของมาโครบล็อกที่ถอดรหัสไม่ได้

เมื่อมีความผิดพลาดในช่องสัญญาณจะทำให้มีบิตผิดพลาดอยู่ในกระแสบิตซึ่งส่งผลกระทบต่อมาโครบล็อกไม่สามารถถอดรหัสกลับคืนได้อย่างถูกต้อง เนื่องจากการเข้ารหัสผลต่างที่ใช้ในตัวเข้ารหัสหรือตัวถอดรหัสวีดิทัศน์มีผลทำให้มาโครบล็อกที่ใช้การทำนายจากมาโครบล็อกที่ผิดพลาดนี้ไม่สามารถถอดรหัสได้ถูกต้องเช่นกัน จากที่กล่าวมาข้างต้นสามารถแบ่งประเภทของมาโครบล็อกที่ถอดรหัสไม่ได้ออกเป็น 2 ชนิด ชนิดที่ 1 คือมาโครบล็อกที่มีบิตผิดพลาดอยู่ในกระแสบิตและชนิดที่ 2 คือ มาโครบล็อกที่ทำนายจากมาโครบล็อกที่มีบิตผิดพลาด

2.3.2.1 มาโครบล็อกที่ถอดรหัสไม่ได้ชนิดที่ 1

การเกิดมาโครบล็อกที่ถอดรหัสไม่ได้ชนิดนี้ ขึ้นตรงอยู่กับอัตราบิตความผิดพลาดของช่องสัญญาณ จากการจำลองการสื่อสารพบว่า ถ้าช่องสัญญาณมีอัตราบิตความผิดพลาดสูงอัตราการเกิดมาโครบล็อกชนิดนี้จะสูงเช่นกัน อย่างไรก็ตามพบว่าอัตราบิตความผิดพลาดของช่องสัญญาณที่เท่ากันแต่มีความยาวของบิตความผิดพลาดต่อเนื่อง (ความยาวเบิรสต์) ต่างกันย่อมมีผลต่อความน่าจะเป็นที่ความผิดพลาดในช่องสัญญาณจะทำให้เกิดมาโครบล็อกที่ถอดรหัสไม่ได้ชนิดนี้ต่างกัน ปริมาณการเกิดมาโครบล็อกชนิดนี้ในการสื่อสารบนช่องสัญญาณไร้สาย ในเฟตติงช้ามีปริมาณการเกิดมาโครบล็อกชนิดนี้น้อยกว่ากว่าการสื่อสารบนช่องสัญญาณไร้สายเฟตติงเร็ว

2.3.2.1 มาโครบล็อกที่ถอดรหัสไม่ได้ชนิดที่ 2

การเกิดมาโครบล็อกที่ถอดรหัสไม่ได้ชนิดนี้ มีการทำนายจากมาโครบล็อกที่มีบิตผิดพลาดชนิดที่ 1 เนื่องจากไม่มีข้อมูลในการทำนายสำหรับถอดรหัสมาโครบล็อก เป็นการสูญเสียมาโครบล็อกที่ขึ้นกับการจัดเรียงมาโครบล็อกในเฟรม หากมาโครบล็อกถูกจัดลำดับไว้อยู่ถัดจากมาโครบล็อกที่ถอดรหัสไม่ได้ชนิดที่ 1 คือมาโครบล็อกเข้ารหัสโดยทำนายจากมาโครบล็อกที่มีบิตผิดพลาด หรือเป็นมาโครบล็อกที่ถูกนำไปใช้ในการทำนายมาโครบล็อกอื่น ก็จะส่งผ่านการถอดรหัสไม่ได้ไปยังมาโครบล็อกต่อไปด้วย จากการจำลองการสื่อสารพบว่าลักษณะการเกิดมาโครบล็อกที่ถอดรหัสไม่ได้

ชนิดที่ 2 นี้ เป็นส่วนใหญ่ของจำนวนมาโครบล็อกที่ถอดรหัสไม่ได้ภาพในเฟรม เนื่องจากลักษณะของการเข้ารหัสที่ต้องการลดจำนวนบิตจึงใช้การเข้ารหัสแบบทำนายเป็นส่วนมาก

2.4 แบบจำลองทางภาษามือ

ภาษามือเป็นอวัจนภาษาอย่างหนึ่งที่ใช้ลักษณะของมือทำเป็นสัญลักษณ์เพื่อใช้ในการสื่อสารท่าของภาษามือจะประกอบไปด้วย 5 หน่วยพื้นฐานย่อย [4]

1. รูปร่างท่ามือ (Hand shape)
2. ตำแหน่ง (Location)
3. ทิศทาง (Orientation)
4. การเคลื่อนไหว (Movement)
5. สัญลักษณ์ที่ไม่ได้ทำด้วยท่ามือ (Non-manual Signal)

จากแบบจำลองทางภาษามือ (Linguistic model) ข้อมูลของภาษามือจะประกอบไปด้วยสองส่วนที่สำคัญคือการแสดงออกทางสีหน้าและท่ามือ การเข้าใจความหมายที่ถูกต้องของการแสดงออกทางสีหน้าถือว่าเป็นสิ่งสำคัญซึ่งการเปลี่ยนแปลงเพียงเล็กน้อยในการแสดงออกทางสีหน้าจะเปลี่ยนความหมายของภาษามืออย่างมาก ส่วนท่ามือหรือการสะกดนิ้วมือ (Finger spelling) จะถูกใช้ในการสะกดคำ ในแง่ของพฤติกรรมของผู้ใช้เมื่อมีการสื่อสารโดยใช้ภาษามือคู่สนทนาจะมีแนวโน้มที่สนใจบริเวณใบหน้าของผู้ทำท่ามือ ในงานวิจัย [12, 13] ศึกษาการติดตามของดวงตา (Eye tracking) โดยแสดงให้เห็นว่าผู้ใช้ภาษามือถึง 95% สนใจที่จะมองในบริเวณหน้ามากกว่าสนใจมองในบริเวณมือ จากลักษณะของการเข้าใจความหมายในภาษามือ หลายงานวิจัยจึงให้ความสำคัญกับบริเวณใบหน้าและมือของผู้ทำท่าภาษามือโดยให้ความสำคัญเล็กน้อยกับบริเวณลำตัวและไม่สนใจบริเวณพื้นหลังของวิดีโอทัศนภาษามือ

2.5 การวัดคุณภาพสัญญาณวิดีโอทัศน

ในการวัดคุณภาพของภาพที่สร้างขึ้นจากกระบวนการชดเชยการเคลื่อนที่ด้วยวิธีประมาณการเคลื่อนที่ ค่าที่นิยมใช้ตรวจสอบว่าภาพที่เข้ารหัสมีความใกล้เคียงกับภาพต้นแบบมากน้อยเพียงใดคือค่าอัตราส่วนสัญญาณยดต่อสัญญาณรบกวนซึ่งเป็นค่าเชิงประวิสัย (Objective) สำหรับวิดีโอทัศนภาษามือคุณภาพของวิดีโอทัศนคือความง่ายต่อการแปลความหมายได้ถูกต้อง ค่าความชัดเจนในการแปลความหมายของวิดีโอทัศนจึงถูกนำมาใช้ในการวัดคุณภาพ ค่าความชัดเจนในการแปลความหมายเป็นค่าเชิงประวิสัย ที่ให้ค่าใกล้เคียงกับค่าเชิงอัตวิสัย (Subjective)

2.5.1 ค่าอัตราส่วนสัญญาณยอดต่อสัญญาณรบกวน (Peak Signal to Noise Ratio: PSNR)

สามารถคำนวณได้ตามสมการที่ (2.27)

$$PSNR = 10 \log_{10} \left[\frac{(2^n - 1)^2}{MSE} \right] \quad (2.27)$$

เมื่อ n คือ จำนวนบิตที่ใช้แทนค่าสีในแต่ละจุดภาพ

MSE คือ ค่าผิดพลาดกำลังสองเฉลี่ยของพิกเซลต้นฉบับกับพิกเซลที่ผ่านการถอดรหัส

2.5.2 ค่าความชัดเจนในการแปลความหมายของวิดิทัศน์ภาษามือ (Computation Intelligibility model: CIM) [5]

ค่าความชัดเจนในการแปลความหมายเป็นอัตราส่วนระหว่างค่าคงที่ของค่าความชัดเจนในการแปลความหมายต่อความเพี้ยนที่มีผลต่อค่าความชัดเจนในการแปลความหมาย โดยให้ความสำคัญกับในแต่ละบริเวณของวิดิทัศน์ที่ไม่เท่ากันตามสมการที่ (2.28)

$$CIM = \log_{10} \left(\frac{110^2}{D_{intell}} \right) \quad (2.28)$$

โดยที่ D_{intell} คือค่าความเพี้ยนในบริเวณใบหน้า มือและลำตัวที่มีผลต่อการแปลความหมายที่ถูกต้องของภาษามือ [14] ตามสมการที่ (2.29)

$$D_{intell} = \alpha_F D_F(n) + \alpha_H D_H(n) + \alpha_T D_T(n) + \alpha_{BG} D_{BG}(n) \quad (2.29)$$

โดยที่ D_F, D_H, D_T และ D_{BG} คือค่าความเพี้ยนในบริเวณใบหน้า มือ ลำตัวและพื้นหลังในเฟรมที่ n และค่าถ่วงน้ำหนักถูกกำหนดให้มีค่า $\alpha_F = 1.6, \alpha_H = 0.5, \alpha_T = 0.1$ และ $\alpha_{BG} = 0$ ตามลำดับ

2.5.3 ค่าสัมประสิทธิ์การตัดสินใจ (Coefficient of Determination : R^2)

เป็นค่าที่บอกถึงความสัมพันธ์ระหว่างตัวแปรทั้ง 2 ว่ามีความสัมพันธ์กันมากน้อยเพียงใด สามารถคำนวณได้ตามสมการที่ (2.30)

$$R^2 = 1 - \frac{\sum_i (Y_i - \hat{Y}_i)^2}{\sum_i (Y_i - \bar{Y}_i)^2} \quad (2.30)$$

โดยที่ Y_i และ \hat{Y}_i คือค่าจริงและค่าที่ประมาณได้ของข้อมูลที่ตำแหน่ง i ตามลำดับและ \bar{Y} คือค่าเฉลี่ย เมื่อค่าสัมประสิทธิ์การตัดสินใจมีค่าใกล้เคียง 1 แสดงว่าค่าที่ได้จากการประมาณมีค่าใกล้เคียงกับค่าจริง

2.6 เอกสารและงานวิจัยที่เกี่ยวข้อง

สำหรับการสื่อสารผ่านโทรศัพท์มือถือเป็นการสื่อสารแบบ 2 ทิศทาง ผลกระทบความความกว้างช่องสัญญาณที่จำกัดจึงขึ้นอยู่กับความเร็วในการเชื่อมโยงขึ้น (uplink) ของสัญญาณ สำหรับเครือข่ายยุค 3G ด้วยเทคโนโลยี HSDPA ให้ความเร็วในการเชื่อมโยงขึ้นของสัญญาณที่ 384 kbps รองรับความละเอียดของวิดีโอที่มีขนาดตั้งแต่ QCIF (176 x 144 จุดภาพ) จนถึง VGA (640 x 480 จุดภาพ) [15, 16] โดยความละเอียดต่ำสุดที่เหมาะสมสำหรับวิดีโอภาษามือถือคือ ขนาด CIF (352 x 288 จุดภาพ) เนื่องจากให้รายละเอียดของมือที่ชัดเจน [17] การเข้ารหัสที่อัตราบิตต่ำคือการเข้ารหัสที่คุณภาพของวิดีโอภาษามือถือยังคงใช้ในการแปลความหมายได้ สำหรับวิดีโอขนาด CIF เมื่อทำการเข้ารหัสที่อัตราบิตมากกว่า 256 kbps จะไม่มีผลต่อความพึงพอใจของผู้ใช้ภาษามือถือในแง่ของคุณภาพวิดีโอที่ 128 kbps ส่งผลต่อการเสื่อมของคุณภาพละเอียดที่อัตราบิตต่ำกว่า 64 kbps ไม่สามารถให้ความพึงพอใจต่อผู้ใช้ภาษามือถือและเพียงพอต่อการเข้ารหัสที่อัตราเฟรมเหมาะสมสำหรับวิดีโอภาษามือถือได้ [15] สำหรับการเข้ารหัสที่อัตราบิตต่ำเพื่อไม่ให้สูญเสียการเคลื่อนที่ที่สำคัญของท่ามือและความชัดเจนในการแปลความหมายในวิดีโอจึงควรมีการเข้ารหัสที่อัตราเฟรมมากกว่า 10 เฟรมต่อวินาที โดยในช่วง 10 – 15 เฟรมต่อวินาทีเป็นช่วงที่ผู้ใช้พึงพอใจ [18]

ในช่วงหลายปีที่ผ่านมาได้มีงานวิจัยที่พัฒนาเทคโนโลยีด้านการสื่อสารผ่านวิดีโอเพื่อช่วยเหลือคนหูหนวกในการติดต่อสื่อสารระหว่างกัน โดยพัฒนาตัวเข้ารหัสด้วยวิธีแบบบล็อกซึ่งให้ประสิทธิภาพในการเข้ารหัสแบบเป็นธรรมชาติสำหรับวิดีโอภาษามือถือ โดยทั่วไปการเข้ารหัสด้วยวิธีแบบบล็อกที่ออกแบบมาสำหรับวิดีโอภาษามือถือจะมีการจัดสรรบิตในอัตราส่วนที่มากกว่าให้กับบล็อกที่สำคัญเช่นบริเวณใบหน้าหรือมือของผู้ทำท่าภาษามือถือ โดยรายละเอียดของวิธีการเข้ารหัสที่ออกแบบสำหรับวิดีโอภาษามือถือสามารถแบ่งได้เป็น 2 ระดับคือ การเข้ารหัสวิดีโอแบบโฟเวียด (foveated video encoding) [15, 19] และการเข้ารหัสในบริเวณที่สนใจ (region of interest encoding) [20, 21] การเข้ารหัสวิดีโอแบบโฟเวียดเป็นการเข้ารหัสที่ไม่มีแบบแผนแน่นอนโดยใช้พฤติกรรมการมองของมนุษย์ สำหรับวิดีโอภาษามือถือผู้ชมสนใจที่จะมองในบริเวณใบหน้าของผู้ทำท่าภาษามือถือจึงกำหนดให้บริเวณนี้เป็นจุดโฟเวียด โดยบริเวณที่อยู่ใกล้จุดโฟเวียดซึ่งถือว่าเป็นบริเวณที่สำคัญจะได้รับการจัดสรรบิตที่มีปริมาณมากกว่าบริเวณที่อยู่ไกลจากจุดโฟเวียด การเข้ารหัสบริเวณ ROI ใช้คุณสมบัติของภาษามือถือที่มีการสื่อสารผ่านทางใบหน้าและมือของผู้ทำท่าภาษามือถือ โดยเข้ารหัสบริเวณที่สนใจด้วยค่าควอนไทซ์พารามิเตอร์ที่ต่ำกว่าจึงส่งผลให้ได้คุณภาพบริเวณใบหน้าและมือสูง

กว่าบริเวณอื่นๆภายในเฟรม เมื่อทำการเข้ารหัสที่อัตราบิตต่ำการเข้ารหัสแบบโพเวียดสามารถลดจำนวนบิตในการเข้ารหัสได้โดยใช้การเพิ่มค่าควอนไทซ์ในบริเวณที่ห่างไกลจากใบหน้าออกไปเรื่อยๆ เนื่องจากการเข้ารหัสแบบนี้ไม่มีการกำหนดความสำคัญของบริเวณผู้ทำท่าภาษามือกับพื้นหลังที่ชัดเจนทำให้พื้นหลังที่อยู่ใกล้บริเวณใบหน้าได้รับการจัดสรรบิตในอัตราส่วนที่สูงตั้งนั้นเพื่อให้ประสบความสำเร็จในการเข้ารหัสที่อัตราบิตต่ำมากๆบริเวณใบหน้าจะถูกควอนไทซ์ด้วยค่าที่สูงขึ้นทำให้เกิดปัญหาภาพเบลอและเป็นวง (blurring and ringing) การเข้ารหัสในบริเวณที่สนใจจะประสบความสำเร็จในการเข้ารหัสที่อัตราบิตต่ำมากกว่าแบบโพเวียดเนื่องจากการแบ่งบริเวณใบหน้าและมือที่ชัดเจนตั้งนั้นที่อัตราบิตต่ำมากๆจึงไม่จำเป็นต้องจัดสรรอัตราบิตให้กับบริเวณพื้นหลัง แต่การเข้ารหัสแบบบริเวณที่สนใจนี้อาศัยวิธีการแยกแยะบริเวณใบหน้าและมือขึ้นสูงซึ่งเป็นการเพิ่มความซับซ้อนให้กับตัวเข้ารหัส

MobileASL Project [22] เป็นอีกกลุ่มที่ทำวิจัยเกี่ยวกับเทคโนโลยีการสื่อสารสำหรับคนหูหนวก โดยพัฒนาระบบโทรศัพท์สำหรับคนพิการโดยใช้ตัวเข้ารหัส x264 ในการเข้ารหัสวีดิทัศน์บนโทรศัพท์เคลื่อนที่เพื่อให้นักพิการใช้ในการติดต่อสื่อสารผ่านทางภาษามือ สิ่งที่ได้นำเสนอในโครงการ MobileASL คือการศึกษาพฤติกรรมของผู้ใช้ภาษามือที่มีต่อการเข้ารหัสวีดิทัศน์เพื่อหาค่าพารามิเตอร์ที่เหมาะสมสำหรับการเข้ารหัสที่ให้ความง่ายต่อการแปลความหมายได้ถูกต้องสูง งานวิจัยในโครงการนี้อาจเช่น ใน [23] สำรวจพฤติกรรมและความเห็นที่มีต่อการเข้ารหัสวีดิทัศน์ของผู้ที่ใช้ภาษามือในการสื่อสารทั้งที่ใช้เป็นภาษาหลักคือคนหูหนวกและที่ใช้เป็นภาษาที่ 2 คือล่ามและคนทั่วไป โดยเป็นการสำรวจความพึงพอใจในด้านต่างๆเช่น การแปลความหมายยากง่ายเพียงใด พอใจกับวีดิทัศน์ที่ผ่านการเข้ารหัสแบบไหนและเปรียบเทียบความสัมพันธ์เชิงปริสัยและเชิงอัติวิสัยเมื่อทำการเข้ารหัสที่ขนาดของวีดิทัศน์และอัตราบิตแตกต่างกัน [1] นำเสนอขั้นตอนวิธี อัตราเฟรมเปลี่ยนแปลงได้ (variable frame rate : VFR) โดยพิจารณาคุณสมบัติของช่วงการสนทนาภาษามือกล่าวคือกำหนดให้ช่วงเฟรมในวีดิทัศน์ขณะกำลังทำท่าภาษามือมีอัตราเฟรมมากกว่าในช่วงอื่น ซึ่งช่วยเพิ่มการใช้พลังงานของแบตเตอรี่อุปกรณ์สื่อสารและเหมาะสมกับความกว้างช่องสัญญาณที่ต่ำโดยไม่ทำให้สูญเสียความชัดเจนในการแปลความหมายของวีดิทัศน์ การใช้อัตราเฟรมเปลี่ยนแปลงได้ก่อให้เกิดการรับรู้วีดิทัศน์ที่เร็ว งานวิจัย [3] จึงนำเสนอขั้นตอนวิธีความละเอียดเชิงพื้นที่เปลี่ยนแปลงได้ (variable spatial resolution : VSR) โดยนำมาใช้ร่วมกับขั้นตอนวิธีอัตราเฟรมเปลี่ยนแปลงได้ ซึ่งมีความคล้ายกันต่างกันตรงที่ในขณะที่กำลังทำท่าภาษามือจะกำหนดให้เฟรมนั้นมีความละเอียดมากกว่าบริเวณอื่นๆ ขั้นตอนวิธีความละเอียดเชิงพื้นที่เปลี่ยนแปลงได้จะส่งผลให้วีดิทัศน์เบลอไม่ชัด ถึงแม้ว่าทั้งสองจะมีผลกระทบต่อคุณภาพของวีดิทัศน์แต่การใช้ทั้งสองขั้นตอนวิธีร่วมกันจะให้คุณภาพการรับรู้ที่ดีกว่าการใช้เพียงแค่วิธีใดวิธีหนึ่ง

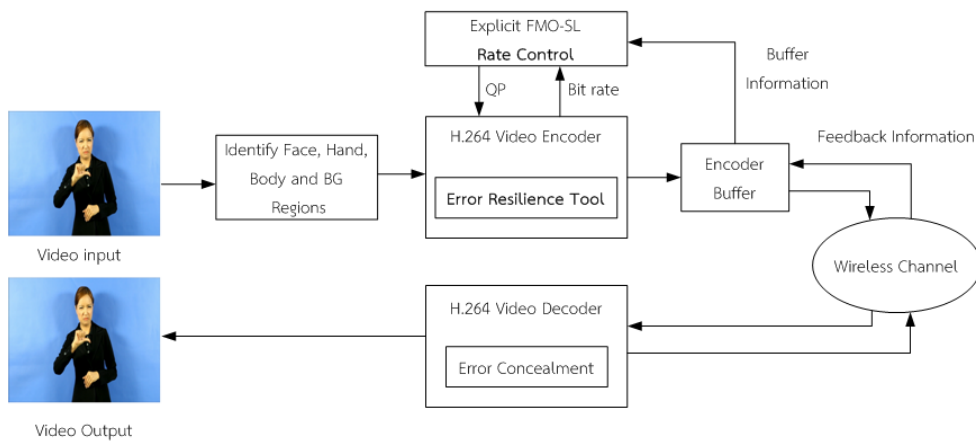
ใน [4, 5, 14, 24] นำเสนอการวัดค่าความเพี้ยนสำหรับวิดิทัศน์ภาษามือโดยให้ความสำคัญในแต่ละบริเวณที่แตกต่างกัน ใน [24] ใช้การวัดค่าความเพี้ยนในการวัดคุณภาพในเชิงความชัดเจนในการแปลความหมายของวิดิทัศน์ภาษามือโดยใช้ข้อมูลเชิงพื้นที่บริเวณใบหน้าและมือในการคำนวณ โดยจากแบบสำรวจผู้ใช้งานพบว่าการวัดคุณภาพด้วยค่า PSNR มีความล้มเหลวในวัดคุณภาพเชิงความชัดเจนในการแปลความหมาย ใน [14] นำเสนอแบบจำลองความเพี้ยนในบริเวณใบหน้า มือ ลำตัวและพื้นหลังเพื่อใช้ในกระบวนการควบคุมอัตรา-ความเพี้ยนที่เหมาะสม โดยมีค่าถ่วงน้ำหนักในแต่ละบริเวณที่แตกต่างกันไปเพื่อให้ได้คุณภาพของวิดิทัศน์ที่เหมาะสมทั้งในเชิงของคุณภาพโดยรวมของวิดิทัศน์และความชัดเจนในการแปลความหมายของวิดิทัศน์ ใน [5] นำเสนอแบบจำลองความเพี้ยนในบริเวณผู้ทำท่าภาษามือเช่นเดียวกับงานวิจัยที่ผ่านมาก่อนหน้านี้แต่มีการเพิ่มการพิจารณาค่าความเพี้ยนเชิงเวลาเพิ่มเติมจากเดิมที่มีการพิจารณาเฉพาะเชิงพื้นที่ นอกจากนี้ยังมีการเพิ่มคุณภาพบริเวณพื้นหลังเพื่อให้คุณภาพโดยรวมของวิดิทัศน์สูงขึ้นแต่จะไม่สูงมากกว่าในบริเวณผู้ทำท่าภาษามือและยังคงรักษาคุณภาพในเชิงความชัดเจนในการแปลความหมายของวิดิทัศน์เมื่อมีอัตราบิดมากเพียงพอ ใน [4] ได้นำเสนอแบบจำลองการคำนวณความชัดเจนในการแปลความหมายของวิดิทัศน์ภาษามือซึ่งให้ความถูกต้องในการวัดคุณภาพมากกว่าแบบจำลองการคำนวณแบบเดิมๆ มีการพิจารณาค่าความเพี้ยนทั้งเชิงพื้นที่และเวลา นอกจากนี้ยังมีการพิจารณาผลกระทบของกระบวนการเข้ารหัสเช่นอัตราเฟรมและโหมดการเข้ารหัสที่มีผลต่อการสูญเสียความเข้าใจง่าย จากงานวิจัยเกี่ยวกับภาษามือที่กล่าวมาข้างต้นจะเห็นได้ว่าส่วนใหญ่สนใจที่จะแก้ปัญหาในเรื่องของการเข้ารหัสวิดิทัศน์ที่อัตราบิดต่ำ ใน [6] นำเสนอการหยุดการแพร่กระจายของความผิดพลาดในเชิงเวลาที่เกิดจากข้อมูลมีการสูญหายในขณะส่งผ่านช่องสัญญาณไร้สาย โดยใช้การเข้ารหัสภายในเฟรม (I-frame) และการใช้ข้อมูลป้อนกลับ (feedback information) เมื่อที่ฝั่งตัวถอดรหัสตรวจพบการสูญหายของเฟรมจะมีการร้องขอให้ฝั่งตัวเข้ารหัสส่ง I-frame มาใหม่ แต่ถ้ามีการร้องขอให้ส่ง I-frame อีกครั้งก่อนที่ฝั่งตัวถอดรหัสจะได้รับ I-frame ในครั้งแรก การร้องขอในครั้งที่ 2 จะถูกยกเลิกและจะใช้ I-frame ที่ได้รับในครั้งแรกแทน แต่การใช้ I-frame จะเป็นการเพิ่มบิตให้มากขึ้นซึ่งไม่เหมาะกับการสื่อสารแบบความกว้างช่องสัญญาณต่ำจึงมีการปรับปรุงการป้อนกลับ I-frame ซึ่งพิจารณาจากความสัมพันธ์เชิงเวลาจากการทำนายระหว่างเฟรม โดยในบริเวณของพื้นหลังไม่จำเป็นต้องใช้โหมดการเข้ารหัสภายในเฟรม (intra mode) เพราะมีการเคลื่อนไหวที่ต่ำ การใช้เพียงแคโหมดการเข้ารหัสแบบกระโดด (skip mode) สามารถปรับปรุงคุณภาพของวิดิทัศน์ได้ จากความสัมพันธ์เชิงเวลานี้จึงเลือกใช้การป้อนกลับ P-frame แทนการป้อนกับ I-frame

การใช้การเข้ารหัสภายในเฟรมหรือที่เรียกว่าการคืนสภาพภายในเฟรม (intra frame refresh) เป็นเครื่องมือที่ส่วนให้ต้านทานต่อความผิดพลาดอย่างมีประสิทธิภาพที่นิยมใช้ในการ

เข้ารหัสวีดิทัศน์ แต่เนื่องด้วยมีการใช้บิตเป็นจำนวนมากในการเข้ารหัสจึงไม่สามารถใช้ได้บ่อยครั้ง เพื่อให้ลดจำนวนบิตที่จะใช้ในการเข้ารหัสจึงมีการเลือกใช้ในระดับมาโครบล็อกเท่านั้นหรือที่เรียกว่า การคืนสภาพมาโครบล็อกภายใน (intra macroblock refresh) ในมาตรฐาน H.264 [25] มีการเลือก มาโครบล็อกแบบสุ่มเพื่อใช้การคืนสภาพมาโครบล็อกภายใน (random intra refresh) ใน [26] นำเสนอการใช้การคืนสภาพมาโครบล็อกภายในด้วยแบบจำลองการสนใจของมนุษย์คือมีการใช้การ คืนสภาพมาโครบล็อกภายในเฉพาะในบริเวณที่มนุษย์ให้ความสนใจ เช่นในบริเวณใบหน้าของคนใน วิดิทัศน์หรือวัตถุขนาดใหญ่ที่มีการเคลื่อนที่ นอกจากนี้ยังมีการนำเสนอแบบจำลองความผิดพลาดเพื่อ ใช้ในการเลือกมาโครบล็อกที่จะใช้การคืนสภาพมาโครบล็อกในในกระบวนการอัตราบิต-ความเพี้ยนที่ เหมาะสม การใช้การเข้ารหัสมาโครบล็อกภายในจะทำให้คุณภาพของวีดิทัศน์ลดลงเนื่องจากมีการใช้ บิตสูงจึงทำให้จำนวนบิตที่จัดสรรในการเข้ารหัสเฟรมถัดไปน้อยลง ใน [27] นำเสนอการเลือก มาโครบล็อกโดยพิจารณาจากค่าใช้จ่ายอัตราบิต-ความเพี้ยน (RDO cost) มาโครบล็อกที่มีอัตราส่วน ค่าใช้ระหว่างโหมดการเข้ารหัสภายในกับโหมดการเข้ารหัสระหว่างเฟรมไม่เกินค่าขีดแบ่งที่กำหนด จะ ถูกเลือกให้ใช้การคืนสภาพมาโครบล็อกภายในซึ่งวิธีนี้จะช่วยให้วีดิทัศน์ต้านทานต่อความผิดพลาดใน ขณะที่มีการเพิ่มค่าใช้จ่ายอัตราบิต-ความเพี้ยนเพียงเล็กน้อย เพื่อเพิ่มประสิทธิภาพการคืนสภาพ มาโครบล็อกภายในสามารถปรับตัวได้ตามสถานะของความผิดพลาดที่เกิดขึ้นในขณะนั้น ในงานวิจัยที่ [28, 29] ได้พิจารณาสภาพของช่องสัญญาณเพื่อใช้ในการหาอัตราของการคืนสภาพมาโครบล็อก ภายใน

บทที่ 3 เทคนิคที่นำเสนอ

วิทยานิพนธ์นี้นำเสนอเครื่องมือด้านทานความผิดพลาดเพื่อลดผลกระทบของวีดิทัศน์จากการสูญหายของข้อมูลเมื่อมีการนำไปใช้งานบนโครงข่ายไร้สายอย่างมีประสิทธิภาพ ได้แก่การคืนสภาพมาโครบล็อกภายในและการจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้โดยอาศัยข้อมูลป้อนกลับจากตัวถอดรหัส รวมถึงการปรับปรุงการควบคุมอัตราระดับเฟรมและหน่วยพื้นฐานย่อยตามแบบจำลองทางภาษาโดยมีการพิจารณาผลกระทบของเครื่องมือที่ช่วยให้ด้านทานความผิดพลาดร่วมด้วยตามรูปที่ 3.1



รูปที่ 3.1 แผนภาพระบบที่นำเสนอ

3.1 แผนการคืนสภาพมาโครบล็อกภายใน (Intra macroblock refresh scheme)

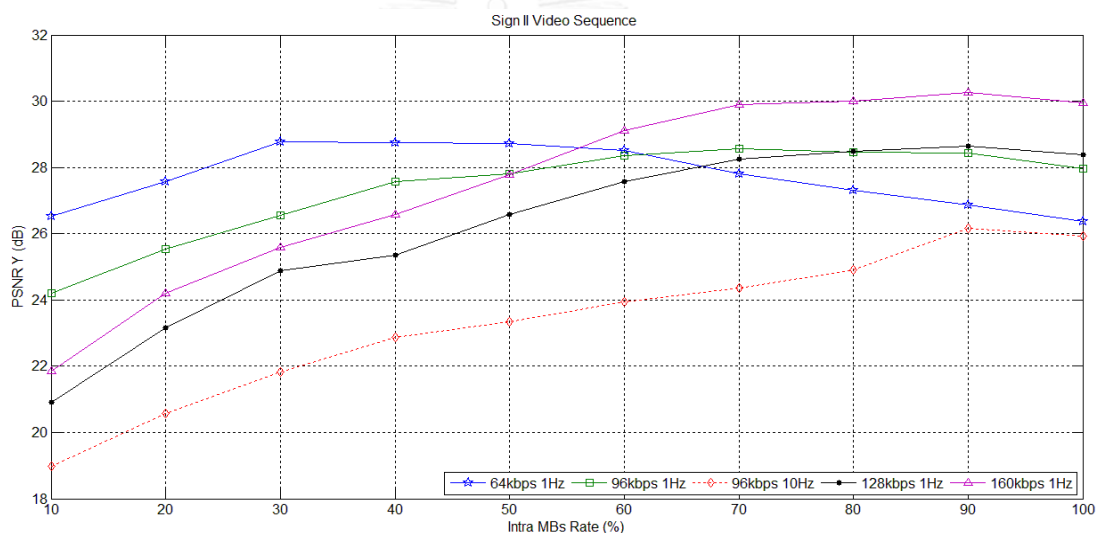
3.1.1 แบบจำลองอัตราคืนสภาพมาโครบล็อกภายใน

การคืนสภาพมาโครบล็อกภายในช่วยในการยับยั้งการแพร่กระจายความผิดพลาดที่เกิดขึ้นภายในเฟรมก่อนหน้ามายังเฟรมที่อยู่ถัดไป อย่างไรก็ตามการคืนสภาพมาโครบล็อกภายในไม่สามารถใช้ได้บ่อยครั้งเนื่องจากการใช้บิตจำนวนมากในการเข้ารหัส การใช้การคืนสภาพมาโครบล็อกภายในที่มากเกินไปจะลดคุณภาพของวีดิทัศน์อย่างมีนัยสำคัญเนื่องจากจะไปลดจำนวนบิตที่จัดสรรสำหรับเข้ารหัสเฟรมถัดไปในกลุ่มภาพ ดังนั้นการใช้อัตราการคืนสภาพมาโครบล็อกภายในที่เหมาะสมภายในอัตราบิตที่ใช้ประโยชน์ได้จึงเป็นประเด็นที่สำคัญ ในวิทยานิพนธ์นี้นำเสนอแบบจำลองอัตราการคืนสภาพมาโครบล็อกภายในตามแบบจำลองทางภาษาโดยให้ความสำคัญกับบริเวณใบหน้าและมือของผู้

ทำท่าภาษามือ เพื่อให้ตัวเข้ารหัสสามารถปรับตัวตามสภาพช่องสัญญาณที่เกิดความผิดพลาดจึงมีการพิจารณาข้อมูลป้อนกลับเพื่อปรับอัตราคืนสภาพมาโครบล็อกภายในอย่างมีประสิทธิภาพ

อัตราคืนสภาพภายในที่เหมาะสมคือการเลือกอัตราร้อยละของจำนวนมาโครบล็อกภายในที่ใช้ในการเข้ารหัสบริเวณใบหน้าและมือต่อจำนวนมาโครบล็อกในบริเวณใบหน้าและมือทั้งหมดในหนึ่งเฟรมที่ทำให้ค่า PSNR เฉลี่ยในวิดิทัศน์สูงที่สุด

การทดลองอย่างละเอียด [28, 29] ถูกนำมาใช้ในการหาแบบจำลองอัตราคืนสภาพภายในที่เหมาะสมรายละเอียดกล่าวไว้ในบทที่ 4 เมื่อทำการทดสอบภายใต้ลำดับภาพ อัตราบิตและการจำลองช่องสัญญาณที่มีความผิดพลาดแตกต่างกัน พบว่าอัตราคืนสภาพมาโครบล็อกภายในที่เหมาะสมที่ให้ค่า PSNR สูงที่สุดนอกจากขึ้นอยู่กับค่าความเพี้ยนที่ได้จากข้อมูลป้อนกลับแล้วยังขึ้นอยู่กับอัตราบิตสำหรับการเข้ารหัสอีกด้วยตามรูปที่ 3.2



รูปที่ 3.2 กราฟ PSNR ของอัตราการคืนสภาพมาโครบล็อกภายในที่แตกต่างกันในแต่ละอัตราบิตและที่ความถี่ดอปเปลอร์ 1 และ 10 Hz ของลำดับภาพ Sign II

ในวิทยานิพนธ์นี้เลือกใช้การเข้ารหัสมาโครบล็อกภายในเฉพาะบริเวณใบหน้าและมือของผู้ทำท่าภาษามือ จึงมีการคำนวณค่าความเพี้ยนเฉพาะในบริเวณใบหน้าและมือเท่านั้น ค่าความเพี้ยนเฉพาะในบริเวณใบหน้าและมือ (D_i) ในเฟรมที่ i^{th} หาได้จากค่าความผิดพลาดเฉลี่ยกำลังสองระหว่างวิดิทัศน์ที่ผ่านการถอดรหัส $f(x, y)$ และวิดิทัศน์ที่ได้จากการสร้างคืน $\tilde{f}(x, y)$ ที่จุดภาพ x, y ของเฟรม $i-1^{\text{th}}$ ตามสมการที่ (3.1) เมื่อ MB_m เป็นมาโครบล็อกตำแหน่งที่ m^{th} ของบริเวณใบหน้าและมือภายใน (FH) ในเฟรม i^{th} ซึ่งมีจำนวนทั้งหมด N_{FH} มาโครบล็อก

$$D_i = \frac{1}{N_{FH}} \cdot \frac{1}{x \cdot y} \cdot \sum_{MB_m \in FH} \sum_{x, y \in MB_m} (f_{i-1}(x, y) - \tilde{f}_{i-1}(x, y))^2 \quad (3.1)$$

ค่ากลุ่มตัวอย่างที่ได้จากการทดลองอย่างละเอียดจะถูกนำมาสร้างเป็นแบบจำลองเพื่อช่วยให้ตัวเข้ารหัสสามารถปรับตัวเพื่อเลือกอัตราคืนสภาพมาโครบล็อกภายในที่เหมาะสมแบบอัตโนมัติได้ในบริเวณใบหน้าและมือต่อเฟรมเมื่อทราบค่าความเพี้ยน (D_i) และบิดเป้าหมายที่ใช้ในการเข้ารหัส (T_i) ในแต่ละเฟรมตามสมการที่ (3.2) โดยค่า % *intra* เป็นร้อยละของจำนวนมาโครบล็อกภายในที่ใช้ในการเข้ารหัสบริเวณใบหน้าและมือต่อจำนวนมาโครบล็อกในบริเวณใบหน้าและมือทั้งหมดในหนึ่งเฟรม

$$\%intra = (a \cdot D_i^2) + (b \cdot T_i) + (c \cdot D_i) + (d \cdot T_i \cdot D_i) \quad (3.2)$$

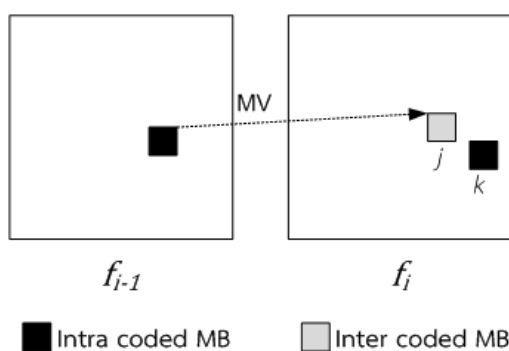
เมื่อ a, b, c, d เป็นค่าพารามิเตอร์ของแบบจำลองที่ได้จากการประมาณโดยใช้ฟังก์ชันไม่เชิงเส้นของชุดข้อมูลที่ได้จากการทดลองอย่างละเอียดด้วยวิธีเลเวนเบิร์ก-มาร์ควอดท์ (the Levenberg-marquardt method) [30] ค่าพารามิเตอร์ที่ได้จากการประมาณในแต่ละลำดับภาพมีความใกล้เคียงกัน ดังนั้นค่าพารามิเตอร์ของแบบจำลองจึงหาได้จากการเฉลี่ยพารามิเตอร์ในแต่ละลำดับภาพ มีค่า $a = -0.391$, $b = 1.13 \times 10^{-4}$, $c = 12.1$ และ $d = 2.43 \times 10^{-5}$

3.1.2 การจัดการคืนสภาพมาโครบล็อกภายในบริเวณใบหน้าและมือของผู้ทำท่าภาษามือ

จำนวนมาโครบล็อกภายในที่ใช้ในการเข้ารหัสบริเวณใบหน้าและมือของผู้ทำท่าภาษามือในวิดิทัศน์ได้จากแบบจำลองอัตราคืนสภาพมาโครบล็อกภายในตามสมการที่ (3.2) โดยมาโครบล็อกที่มีค่าความเพี้ยนสูงจะถูกเปลี่ยนให้เป็นมาโครบล็อกภายในตามจำนวนของมาโครบล็อกที่คำนวณได้จากแบบจำลอง ซึ่งค่าความเพี้ยนของแต่ละมาโครบล็อกในบริเวณใบหน้าและมือของผู้ทำท่าภาษามือสามารถคำนวณได้จากสมการที่ (3.1)

วิธีการคืนสภาพมาโครบล็อกภายในโดยทั่วไปจะเป็นการเพิ่มการใช้มาโครบล็อกภายในไปยังมาโครบล็อกที่มีความซับซ้อนสูงหรือมีความผิดพลาดสูงซึ่งเป็นมาโครบล็อกที่ยากต่อการปกปิดความผิดพลาด ปัญหาของวิธีการคืนสภาพมาโครบล็อกภายในคือไม่สามารถรับประกันได้ว่าทุกมาโครบล็อกในวิดิทัศน์จะได้รับการคืนสภาพ มาโครบล็อกที่ไม่ได้รับการคืนสภาพจะนำไปสู่การแพร่กระจายความผิดพลาดอย่างไม่มีกำหนดในตำแหน่งของมาโครบล็อกเหล่านี้ ในวิทยานิพนธ์นี้ใช้การเข้ารหัสภายในเฉพาะมาโครบล็อกบริเวณใบหน้าและมือของผู้ทำท่าภาษามือโดยพิจารณาจากค่าความเพี้ยนที่เกิดขึ้นในแต่ละมาโครบล็อก โดยมาโครบล็อกที่มีความเพี้ยนมากจะได้รับการคืนสภาพภายในบ่อยครั้งมากกว่ามาโครบล็อกที่มีความเพี้ยนน้อยซึ่งแทบจะไม่ได้ได้รับการคืนสภาพภายใน เพื่อจัดการกับปัญหาการแพร่กระจายความผิดพลาด ใน [28] ใช้การเพิ่มอัตราการคืนสภาพมาโครบล็อกภายในเพื่อให้ทุกมา

โครบล็อคได้รับการคืนสภาพแต่วิธีนี้จะเป็นการเพิ่มการใช้อัตราบิตที่มากขึ้นในการเข้ารหัส ในวิทยานิพนธ์นี้นำเสนอวิธีโดยพิจารณาจากการทำนายการเคลื่อนที่ในช่วงของขั้นตอนการประมาณการเคลื่อนที่ตามรูปที่ 3.3 ก่อนที่มาโครบล็อค j^{th} ในเฟรมปัจจุบันจะถูกเปลี่ยนไปใช้โหมดการเข้ารหัสภายใน มาโครบล็อค j^{th} นี้มีเป็นมาโครบล็อคที่ใช้โหมดการเข้ารหัสระหว่างเฟรม จากนั้นถ้ามาโครบล็อค j^{th} ในเฟรมปัจจุบัน มีการทำนายจากบริเวณมาโครบล็อคที่ใช้โหมดการเข้ารหัสภายในของเฟรมก่อนหน้า มาโครบล็อค j^{th} ในเฟรมปัจจุบันจะไม่ถูกเปลี่ยนให้เป็นมาโครบล็อคที่ใช้โหมดการเข้ารหัสภายใน อย่างไรก็ตามมาโครบล็อค k^{th} ซึ่งมีค่าความเพี้ยนน้อยกว่าในลำดับถัดไปของรายการจัดเรียง จะถูกเปลี่ยนเป็นใช้โหมดการเข้ารหัสภายในแทน วิธีนี้จะช่วยให้มาโครบล็อคในบริเวณใบหน้าและมือของผู้ทำท่าภาษามือได้รับการคืนสภาพ



รูปที่ 3.3 ขั้นตอนการทำนายการเคลื่อนที่

3.1.3 การจัดสรรการคืนสภาพมาโครบล็อคภายในบริเวณลำตัวของผู้ทำท่าภาษามือ

วิดิทัศน์ภาษามือที่ผ่านการถอดรหัสสมักจะมีส่วนของมือและพื้นหลังตกค้างอยู่ในบริเวณลำตัวของผู้ทำท่าภาษามือเนื่องจากผลกระทบของการใช้การปกปิดความผิดพลาดแบบไม่มีการชดเชยการเคลื่อนที่หรือความเพี้ยนที่บริเวณมือแพร่กระจายไปยังบริเวณลำตัว ผลกระทบเหล่านี้เป็นการสร้างความเพี้ยนให้กับหน่วยพื้นฐานย่อยทั้ง 5 ของท่ามือโดยเฉพาะรูปร่างและตำแหน่งของท่ามือซึ่งเป็นการลดความชัดเจนในการแปลความหมายของวิดิทัศน์ภาษามือ ดังนั้นคุณภาพวิดิทัศน์ในบริเวณลำตัวของผู้ทำท่าภาษามือจึงควรได้รับการปรับปรุงให้ดีขึ้น เพื่อเป็นการเพิ่มความชัดเจนในการแปลความหมายของวิดิทัศน์ภาษามือ ในวิทยานิพนธ์นี้เสนอวิธีการใช้การคืนสภาพมาโครบล็อคภายในบริเวณลำตัวของผู้ทำท่าภาษามือสามารถอธิบายขั้นตอนได้ดังนี้

$IF(MB_m \in \text{Body in } f_i \text{ and } MB_m \in \text{Hand in } f_{i-1}) \parallel (MB_m \in \text{Body in } f_i \text{ and } MB_m \in \text{BG in } f_{i-1})$

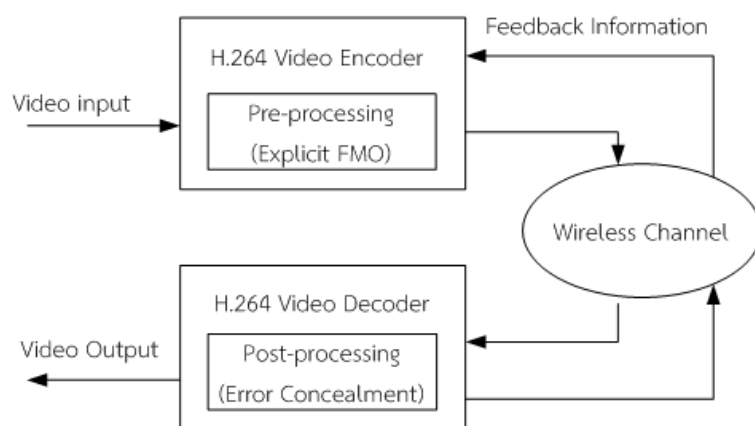
MB_m are forced to intra coding mode

END

เมื่อ MB_m เป็นตำแหน่งมาโครบล็อกที่ m^{th} ของบริเวณมือ (*Hand*) ลำตัว (*Body*) และพื้นหลัง (*BG*) ภายในเฟรม f_i และ f_{i-1}

3.2 การจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้ (Flexible Macroblock Ordering: FMO)

การคืนสภาพมาโครบล็อกภายในเป็นวิธีที่ช่วยยับยั้งการแพร่กระจายความผิดพลาดในเฟรมก่อนหน้าไปยังเฟรมที่อยู่ถัดไป อย่างไรก็ตามแม้ว่าจะมีการยับยั้งการแพร่กระจายความผิดพลาดระหว่างเฟรมแต่ก็ยังมีเกิดการเกิดการแพร่กระจายความผิดพลาดภายในเฟรมขึ้น เนื่องจากการเข้ารหัสผลต่างที่ใช้ในตัวเข้ารหัสมีผลให้มาโครบล็อกที่ใช้การทำนายจากมาโครบล็อกที่ผิดพลาดนี้ไม่สามารถถอดรหัสได้ การจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้คือการเข้ารหัสโดยใช้จำนวนกลุ่มสไลซ์มากกว่า 1 กลุ่มแต่ไม่เกิน 8 กลุ่มสไลซ์ต่อเฟรม หากมีการสูญเสียกลุ่มข้อมูลใดไป การถอดรหัสกลุ่มข้อมูลอื่นก็ยังคงทำได้เนื่องจากไม่มีการทำนายข้ามกลุ่มสไลซ์ การจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้จึงถูกนำมาใช้ในการลดความผิดพลาดชนิดนี้ เนื้อหาในส่วนนี้เป็นการอธิบายเทคนิคการสร้างแผนที่การจัดสรรมาโครบล็อกแบบชัดแจ้งเพื่อใช้ในการจัดกลุ่มของมาโครบล็อกเข้าสู่กลุ่มสไลซ์ แผนที่ถูกสร้างโดยใช้ข้อมูลค่าความเพี้ยนจากการปกปิดความผิดพลาดที่ได้จากข้อมูลป้อนกลับจากตัวถอดรหัสตามรูปที่ 3.4



รูปที่ 3.4 แผนภาพการสร้างแผนที่การจัดสรรมาโครบล็อกแบบชัดแจ้ง

ค่าความเพี้ยน (Dis_{MB}) ถูกใช้เป็นตัวบ่งชี้ความสำคัญของมาโครบล็อกโดยขึ้นอยู่กับประเภทของการปกปิดความผิดพลาดที่ใช้ที่ตัวถอดรหัสคือเมื่อมีความผิดพลาดเกิดขึ้นการปกปิดความ

ผิดพลาดแบบไม่มีการชดเชยการเคลื่อนที่จะถูกนำมาใช้เพื่อปกปิดความผิดพลาด ในงานนี้ค่าความเพี้ยนที่เกิดขึ้นหลังจากการปกปิดความผิดพลาดสามารถคำนวณได้ตามสมการที่ (3.3)

$$Dis_{MB} = \sum_{(x,y \in MB)} |\tilde{f}_{i-1}(x,y) - f_{i-1}(x,y)| \quad (3.3)$$

เมื่อ $\tilde{f}_{i-1}(x,y)$ คือค่าจุดภาพที่ผ่านการปกปิดความผิดพลาดของวิดีโอที่ถอดรหัสที่จุดภาพ x,y ของเฟรมก่อนหน้า $f_{i-1}(x,y)$ คือค่าจุดภาพของวิดีโอต้นฉบับที่จุดภาพ x,y ของเฟรมก่อนหน้าและ MB คือมาโครบล็อกที่มีการปกปิดความผิดพลาดจากความผิดพลาดในช่องสัญญาณ

ขั้นตอนถัดไปในการพัฒนาการเข้ารหัสให้ด้านทานความผิดพลาดโดยใช้การจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้คือการพัฒนาระบบการจัดสรรมาโครบล็อกเข้าสู่กลุ่มสไลซ์นอกเหนือจากวิธีการจัดเรียงแบบแทรกสลับเดิมเพื่อปรับปรุงประสิทธิภาพการเข้ารหัส แนวคิดหลักในการจัดสรรมาโครบล็อกเข้าสู่กลุ่มสไลซ์ที่นำเสนอในวิทยานิพนธ์นี้มีหลักการคล้ายคลึงกับการจัดสรรมาโครบล็อกเข้าสู่กลุ่มสไลซ์ใน [31] ที่มีการจัดมาโครบล็อกเข้าสู่กลุ่มสไลซ์โดยจัดให้แต่ละกลุ่มสไลซ์มีจำนวนบิตที่ใช้เข้ารหัสหรือความเพี้ยนจากการปกปิดความผิดพลาดเท่ากันหรือใกล้เคียงกันมากที่สุด ซึ่งวิธีนี้จะเป็นการพิจารณาค่าความเพี้ยนของทั้งเฟรมโดยไม่ได้พิจารณาเจาะจงว่าค่าความเพี้ยนของมาโครบล็อกบริเวณหน้า มือและลำตัวในแต่ละสไลซ์มีความเท่ากันหรือไม่ ในวิทยานิพนธ์นี้จึงนำเสนอระบบการจัดสรรมาโครบล็อกเข้าสู่กลุ่มสไลซ์อย่างง่ายโดยมีการแบ่งกลุ่มของมาโครบล็อกออกเป็น 2 ประเภทคือกลุ่มของมาโครบล็อกหน้า มือและลำตัว กับกลุ่มของมาโครบล็อกของพื้นหลัง โดยกลุ่มของมาโครบล็อกแรกซึ่งคือกลุ่มของมาโครบล็อกหน้า มือและลำตัวจะถูกจัดเข้าสู่กลุ่มสไลซ์ก่อนแล้วจึงค่อยจัดกลุ่มมาโครบล็อกของพื้นหลังเข้าสู่กลุ่มสไลซ์ ขั้นตอนวิธีการจัดสรรมาโครบล็อกเข้าสู่กลุ่มสไลซ์มีดังนี้

1. จัดเรียงลำดับค่าความเพี้ยนของมาโครบล็อกบริเวณหน้า มือและลำตัวจากมากไปน้อย โดยมาโครบล็อกที่มีค่าความเพี้ยนสูงจะอยู่ในลำดับต้นของรายการจัดเรียงและมาโครบล็อกที่มีค่าความเพี้ยนต่ำจะอยู่ในลำดับล่างของรายการจัดเรียง
2. จัดมาโครบล็อก 2 ลำดับแรกในรายการจัดเรียงไปยังกลุ่มสไลซ์ A และกลุ่มสไลซ์ B จากนั้นนำ 2 มาโครบล็อกแรกออกจากรายการการจัดเรียง
3. คำนวณค่าความผิดพลาดรวมของกลุ่มสไลซ์ A และกลุ่มสไลซ์ B กำหนดให้มีค่าเท่ากับ $SumGrpA$ และ $SumGrpB$ ตามลำดับ
4. จัดมาโครบล็อกปัจจุบัน ($curMB$) ซึ่งอยู่บนสุดของรายการจัดเรียงเข้าสู่กลุ่มสไลซ์ A หรือกลุ่มสไลซ์ B ตามวิธีในข้อถัดไป

5. **IF ($SumGrpA < SumGrpB$)**

curMB ถูกจัดให้อยู่กลุ่มสไลซ์ A

ELSE

curMB ถูกจัดให้อยู่กลุ่มสไลซ์ B

6. นำมาโครบล็อกปัจจุบันออกจากรายการการจัดเรียง

7. ทำซ้ำข้อ 3 ถึง 6 จนกระทั่งทุกมาโครบล็อกของบริเวณหน้า มีและลำตัวได้รับการจัดสรรเข้าสู่กลุ่มสไลซ์

8. ทำซ้ำข้อ 1 ถึง 6 โดยการเปลี่ยนเป็นการจัดเรียงกลุ่มมาโครบล็อกของพื้นหลัง

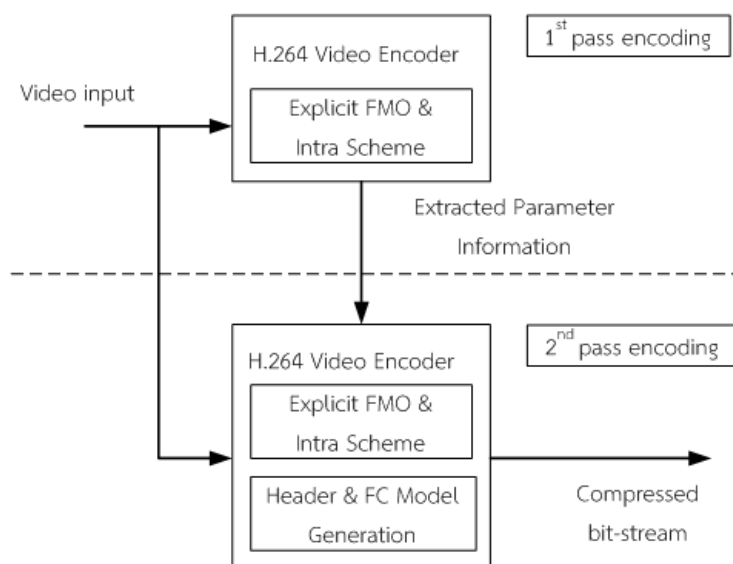
การปรับปรุงกระบวนการจัดสรรมาโครบล็อกเข้าสู่กลุ่มสไลซ์นี้ทำให้สามารถแบ่งมาโครบล็อกออกได้เป็น 2 กลุ่มสไลซ์ที่มีค่าความเพี้ยนรวมของแต่ละกลุ่มสไลซ์ใกล้เคียงกัน นอกจากนี้สามารถใช้การจัดสรรมาโครบล็อกเข้าสู่กลุ่มสไลซ์กับกลุ่มสไลซ์ A และกลุ่มสไลซ์ B ซึ่งจะทำให้แบ่งออกได้เป็นกลุ่มสไลซ์ทั้งหมด 4 กลุ่มสไลซ์และทำซ้ำอีกครั้งจะสามารถแบ่งออกเป็น 8 กลุ่มสไลซ์

3.3 การปรับปรุงการควบคุมอัตราการระดับเฟรมและมาโครบล็อกสำหรับมาตรฐานการเข้ารหัส H.264/AVC

3.3.1 การปรับปรุงการควบคุมอัตราการระดับเฟรมสำหรับมาตรฐานการเข้ารหัส H.264/AVC

การใช้การจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้ช่วยให้วิดิทัศน์ที่ถูกบีบอัดด้านทานความผิดพลาดอย่างมีประสิทธิภาพ ในทางกลับกันจะลดประสิทธิภาพของการเข้ารหัสเนื่องจากจำกัดให้มีการทำนายการเคลื่อนที่เกิดขึ้นเฉพาะในกลุ่มสไลซ์เดียวกันเท่านั้นและเป็นการเพิ่มจำนวนของบิตโดยรวมที่ใช้สำหรับการเข้ารหัสวิดิทัศน์โดยตำแหน่งของมาโครบล็อกในแผนที่การจัดสรรมาโครบล็อกจะถูกระบุลงในบิตส่วนหัว ที่อัตราบิตต่ำบิตส่วนหัวครอบครองสัดส่วนส่วนมากของบิตทั้งหมดเมื่อเทียบกับบิตที่จัดสรรสำหรับเข้ารหัสข้อมูลภาพ ดังนั้นเพื่อให้ได้บิตเป้าหมายในแต่ละเฟรมตรงตามที่ต้องการตัวเข้ารหัสกำหนดการเข้ารหัสวิดิทัศน์จึงควรพิจารณาผลกระทบของบิตส่วนหัวที่เกิดขึ้นเมื่อมีการใช้การจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้ ในวิทยานิพนธ์นี้ปรับปรุงคุณภาพของการควบคุมอัตราการระดับเฟรมของมาตรฐานการเข้ารหัส H.264/AVC โดยนำเสนอแบบจำลองบิตส่วนหัวและปรับปรุงการประมาณบิตเป้าหมายให้มีความสอดคล้องกับความซับซ้อนภายในเฟรมเมื่อการจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้และการคืนสภาพมาโครบล็อกภายในถูกใช้งาน แบบจำลองบิตส่วนหัวและแบบจำลอง

ความซับซ้อนภายในเฟรมถูกสร้างขึ้นโดยใช้การเข้ารหัส 2 รอบ ในรอบแรกเป็นการเข้ารหัสโดยใช้เครื่องมือต้านทานความผิดพลาดเพื่อเก็บค่าพารามิเตอร์ต่างๆที่ใช้ในการคำนวณแบบจำลองทั้ง 2 และรอบที่ 2 เป็นการเข้ารหัสโดยใช้แบบจำลองที่สร้างขึ้นจากข้อมูลที่ได้การเข้ารหัสในรอบแรก ตามที่แสดงในรูปที่ 3.5



รูปที่ 3.5 แผนภาพการสร้างแบบจำลองบิตส่วนหัวและความซับซ้อนภายในเฟรม

1. แบบจำลองบิตส่วนหัว

โดยทั่วไปเวกเตอร์การเคลื่อนที่ของแต่ละมาโครบล็อกที่อยู่ข้างเคียงมักจะมีความสัมพันธ์กัน เนื่องจากการทำนายการเคลื่อนที่ของตัวเข้ารหัสจะใช้การทำนายจากทั้งเฟรม ในมาตรฐาน H.264 ความสัมพันธ์นี้ถูกนำมาใช้ในการทำนายเวกเตอร์การเคลื่อนที่ของมาโครบล็อก โดยตำแหน่งของมาโครบล็อกข้างเคียงที่ใช้ในการทำนายสามารถทราบได้อยู่แล้วสำหรับลำดับการกราดแบบบราสเตอร์ ความแตกต่างของเวกเตอร์การเคลื่อนที่ระหว่างเวกเตอร์จริงและเวกเตอร์ที่ได้จากการทำนายของมาโครบล็อกปัจจุบันจะถูกเข้ารหัสและส่งออกไป ในกรณีที่ใช้การจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้ ลำดับการเข้ารหัสจะถูกเปลี่ยนไป โดยมาโครบล็อกข้างเคียงจะไม่สามารถใช้ได้สำหรับการทำนายการเคลื่อนที่ถ้ามาโครบล็อกเหล่านั้นยังไม่ได้ถูกเข้ารหัสก่อนมาโครบล็อกปัจจุบัน ซึ่งจะส่งผลกระทบต่อ การคำนวณค่าความแตกต่างของเวกเตอร์การเคลื่อนที่และประสิทธิภาพของการลงรหัส ใน [32] วิเคราะห์และให้ความสำคัญกับความแตกต่างของเวกเตอร์การเคลื่อนที่และจำนวนของกลุ่มสไลซ์เพื่อ สร้างแบบจำลองบิตส่วนหัวเมื่อมีการใช้การจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้ โดยแสดงให้เห็นถึงความสอดคล้องกันระหว่างจำนวนกลุ่มสไลซ์และจำนวนความแตกต่างของเวกเตอร์การเคลื่อนที่ที่ไม่

เป็นศูนย์กล่าวคือ เมื่อมีการใช้จำนวนกลุ่มสไลซ์มากขึ้นจำนวนความแตกต่างของเวกเตอร์การเคลื่อนที่ที่ไม่เป็นศูนย์มีจำนวนเพิ่มขึ้นตามไปด้วย ดังนั้นจำนวนความแตกต่างของเวกเตอร์การเคลื่อนที่ที่ไม่เป็นศูนย์ภายในเฟรมสามารถนำมาใช้ในการสร้างแบบจำลองบิตส่วนหัวสำหรับ P เฟรมจำนวนบิตส่วนหัวสามารถคำนวณได้ตามสมการที่ (3.4)

$$H_{\text{pframe}} = \alpha_1 \cdot N_{\text{nzMVD}} + \alpha_2 \cdot (N_{\text{MV}} + \text{num_slice}) \quad (3.4)$$

เมื่อ N_{nzMVD} คือ จำนวนความแตกต่างของเวกเตอร์การเคลื่อนที่ที่ไม่เป็นศูนย์

N_{MV} คือ จำนวนเวกเตอร์การเคลื่อนที่

num_slice คือ จำนวนกลุ่มสไลซ์ที่ใช้ในแต่ละเฟรม

α_1 และ α_2 คือ พารามิเตอร์แบบจำลอง

โดยไม่ได้มีการพิจารณาบิตส่วนหัวของมาโครบล็อกที่เกิดขึ้นจากการเข้ารหัสโดยใช้โหมดการเข้ารหัสภายในเนื่องจากมีจำนวนมาโครบล็อกเหล่านี้อยู่น้อยมากในเฟรมทำให้ไม่มีผลต่อความถูกต้องของแบบจำลอง แต่เนื่องจากในวิทยานิพนธ์นี้มีการใช้การคืนสภาพมาโครบล็อกภายในซึ่งทำให้เกิดมาโครบล็อกที่ใช้โหมดการเข้ารหัสภายในเกิดขึ้น แบบจำลองใน [32] จึงมีผลต่อการประมาณบิตส่วนหัว

ใน [33] นำเสนอแบบจำลองบิตส่วนหัวโดยใช้เวกเตอร์การเคลื่อนที่ในการสร้างแบบจำลองบิตส่วนหัวทั้งหมดภายในเฟรม เป็นการคำนวณบิตส่วนหัวของมาโครบล็อกที่ใช้โหมดการเข้ารหัสระหว่างเฟรมและโหมดการเข้ารหัสภายในเฟรม สามารถคำนวณได้ตามสมการที่ (3.5) โดยบิตส่วนหัวสำหรับการเข้ารหัสภายในเฟรมสามารถคำนวณได้โดยใช้จำนวนของมาโครบล็อกที่ใช้โหมดการเข้ารหัสภายในของเฟรมปัจจุบัน (N_{intra}) และค่าเฉลี่ยบิตส่วนหัวของมาโครบล็อกที่ใช้โหมดการเข้ารหัสภายในในเฟรมก่อนหน้า (b_{intra})

$$R_{\text{hdr}} = \gamma \cdot (N_{\text{nzMVE}} + \omega \cdot N_{\text{MV}}) + N_{\text{intra}} \cdot b_{\text{intra}} \quad (3.5)$$

เมื่อ N_{MV} คือ จำนวนของเวกเตอร์การเคลื่อนที่

N_{nzMVE} คือ จำนวนของเวกเตอร์การเคลื่อนที่ที่ไม่เป็นศูนย์

N_{intra} คือ จำนวนของมาโครบล็อกภายในของเฟรมปัจจุบัน

b_{intra} คือ ค่าเฉลี่ยบิตส่วนหัวของมาโครบล็อกภายในในเฟรมก่อนหน้า

γ คือ ค่าพารามิเตอร์ของแบบจำลอง

ω คือ ค่าถ่วงน้ำหนัก

ในวิทยานิพนธ์นี้ได้ปรับแบบจำลองในสมการที่ (3.4) และสมการที่ (3.5) เพื่อสร้างแบบจำลองบิตส่วนหัวของ P เฟรมเมื่อมีการใช้การจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้ร่วมกับการคืนสภาพมาโครบล็อกภายใน แบบจำลองบิตส่วนหัวของ P เฟรม ($H_{pframe,i}$) ในเฟรมที่ i^{th} สามารถคำนวณได้ตามสมการที่ (3.6)

$$H_{pframe,i} = \alpha_1 \cdot N_{nzMVD} + \alpha_2 (N_{MV} + num_slice) + \alpha_3 (N_{intra} \cdot b_{intra}) \quad (3.6)$$

เมื่อ N_{nzMVD} คือ จำนวนความแตกต่างของเวกเตอร์การเคลื่อนที่ที่ไม่เป็นศูนย์

N_{MV} คือ จำนวนเวกเตอร์การเคลื่อนที่

num_slice คือ จำนวนกลุ่มสไลซ์ที่ใช้ในแต่ละเฟรม

N_{intra} คือ จำนวนของมาโครบล็อกภายในของเฟรมปัจจุบัน

b_{intra} คือ ค่าเฉลี่ยของบิตส่วนหัวของมาโครบล็อกภายในในเฟรมก่อนหน้า

พารามิเตอร์แบบจำลอง $\alpha = [\alpha_1, \alpha_2, \alpha_3]^T$ เป็นค่าที่ทำให้เกิดความผิดพลาดน้อยที่สุดระหว่างปริมาณบิตส่วนหัวที่ได้จากแบบจำลอง (H) ซึ่งเป็นค่าที่ต้องการกับปริมาณบิตส่วนหัวจริงที่ได้จากการทดลองเมื่อมีการใช้การจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้ จากสมการที่ (3.6) สามารถเขียนอยู่ในรูปฟังก์ชันการถดถอยเชิงเส้น (linear regression) ได้ตามสมการที่ (3.7)

$$H = NN^T \alpha \quad (3.7)$$

เมื่อ $H = [H_1, H_2, H_3]^T$ เป็นปริมาณของบิตส่วนหัวที่ได้จากการทดลอง โดยที่

$H_1 = H_{pframe,i} \cdot N_{nzMVD,i}$, $H_2 = H_{pframe,i} \cdot N_{MV,i} + Slice$ และ $H_3 = H_{pframe,i} \cdot N_{intra,i} \cdot b_{intra,i}$ โดย

$N = [N_{nzMVD}, N_{MV} + Slice, N_{intra} \cdot b_{intra}]^T$ ซึ่งการถดถอยเชิงเส้นหาได้จากผลรวมของค่าภายใน

$$\text{เฟรม} \sum_{i=1}^N H_i = \sum N_i N_i^T \alpha$$

2. การวัดความซับซ้อนภายในเฟรม

ในวิทยานิพนธ์นี้ปรับปรุงการควบคุมอัตราการระดับเฟรมให้ดีขึ้นและเหมาะสมสำหรับวิดีโอภาษาเมื่อโดยปรับปรุงการประมาณบิตเป้าหมายให้มีความสอดคล้องกับความซับซ้อนภายในเฟรม การปรับปรุงนี้ใช้แบบจำลองความซับซ้อนภายในเฟรมจาก [32] โดยใช้จำนวนของเวกเตอร์การเคลื่อนที่ซึ่งมีความสัมพันธ์กับความซับซ้อนภายในเฟรมและจำนวนของบิตที่ใช้ในการเข้ารหัส การวัดความซับซ้อนภายในเฟรม (FC_i) ของเฟรมที่ i^{th} สามารถคำนวณได้จากค่าอัตราส่วนผลต่างสัมบูรณ์

($MAD_{ration,i}$) ตามสมการที่ (3.8) และอัตราส่วนจำนวนผลต่างของเวกเตอร์การเคลื่อนที่ที่ไม่เป็นศูนย์ ($N_{nzMVDration,i}$) ตามสมการที่ (3.9)

$$MAD_{ration,i} = \frac{MADP_i}{\frac{1}{(i-1)} \sum_{j=1}^{i-1} MADP_j} \quad (3.8)$$

เมื่อ $MADP_i$ คือ ค่าผลต่างสัมบูรณ์เฉลี่ยที่ทำนายได้ในเฟรม i^{th}

$$N_{nzMVDration,i} = \frac{N_{nzMVD,i}}{\frac{1}{(i-1)} \sum_{j=1}^{i-1} N_{nzMVD,i}} \quad (3.9)$$

เมื่อ $N_{nzMVD,i}$ คือ จำนวนผลต่างของเวกเตอร์การเคลื่อนที่ที่ไม่เป็นศูนย์ในเฟรม i^{th}

การวัดความซับซ้อนภายในเฟรม i^{th} สามารถคำนวณตามสมการ (3.10) โดยที่ค่าพารามิเตอร์ของแบบจำลอง β มีค่าเท่ากับ 0.6 สำหรับวิดิทัศน์ภาษามือที่มีความซับซ้อนสูงและ 0.4 สำหรับวิดิทัศน์ภาษามือทั่วไป

$$FC_i = \beta \cdot MAD_{ration,i} + (1-\beta) \cdot N_{nzMVDration,i} \quad (3.10)$$

จำนวนบิตเป้าหมายทั้งหมดที่ถูกจัดสรรสำหรับเฟรมที่ i^{th} ถูกกำหนดโดยขึ้นอยู่กับบิตเป้าหมายที่คำนวณจากสถานะของบัฟเฟอร์และบิตเป้าหมายที่คำนวณจากจำนวนบิตที่เหลืออยู่สำหรับการเข้ารหัส

บิตเป้าหมายที่คำนวณจากสถานะบัฟเฟอร์ ($T_{buf,i}$) สำหรับเฟรมที่ i^{th} คำนวณได้จากบิตสะสมในบัฟเฟอร์ปัจจุบัน (CBF) ระดับบัฟเฟอร์เป้าหมาย (TBL) อัตราเฟรม (f_r) และอัตราบิต (b_r) ที่สามารถใช้ได้ ตามสมการที่ (3.11) โดย Γ เป็นค่าคงที่ที่กำหนดให้มีค่าเท่ากับ 0.5

$$T_{buf,i} = \left[\frac{b_r}{f_r} - \Gamma(CBF_{i-1} - TBL_i) \right] \quad (3.11)$$

และบิตเป้าหมาย ($T_{rem,i}$) ที่คำนวณจากจำนวนบิตที่เหลืออยู่สำหรับการเข้ารหัสสำหรับเฟรมที่ i^{th} ตามสมการที่ (3.12)

$$T_{rem,i} = \frac{R_i}{N_i} \quad (3.12)$$

โดยที่ R_i คือจำนวนของบิตที่เหลืออยู่ในกลุ่มภาพ และ N_i คือ จำนวนของ P เฟรมที่ไม่ได้ทำการเข้ารหัส

เพื่อให้การประมาณบิตเป้าหมายถูกต้องมากขึ้น การหาบิตเป้าหมายจึงมีการพิจารณาร่วมกับความซับซ้อนของเฟรม บิตเป้าหมาย (T_{mod}) ที่มีการปรับปรุงสามารถคำนวณตามสมการที่ (3.13)

$$T_{\text{mod},i} = \begin{cases} FC_i \cdot T_{\text{rem},i} & ; 0 < FC_i < 1.0 \\ 1.1 \cdot T_{\text{rem},i} & ; 1.0 \leq FC_i < 1.2 \\ 1.2 \cdot T_{\text{rem},i} & ; 1.2 \leq FC_i \end{cases} \quad (3.13)$$

จำนวนของบิต (T_i) ทั้งหมดที่จัดสรรสำหรับเฟรมที่ i^{th} สามารถคำนวณได้จากการถ่วงน้ำหนักของบิตเป้าหมายทั้งสองร่วมกันตามสมการที่ (3.14)

$$T_i = \beta_r \cdot T_{\text{mod},i} + (1 - \beta_r) \cdot T_{\text{buf},i} \quad (3.14)$$

หลังจากคำนวณบิตเป้าหมายแล้ว จำนวนของบิตที่จัดสรรสำหรับข้อมูลภาพ จะถูกคำนวณโดยนำจำนวนของบิตส่วนหัวที่คำนวณได้จากแบบจำลองตามสมการที่ (3.6) มาหักลบออกจากบิตเป้าหมาย (T_i) ตามสมการที่ (3.15)

$$T_{\text{xt},i} = T_i - H_{\text{Pframe},i} \quad (3.15)$$

3.3.2 การปรับปรุงการควบคุมอัตราระดับหน่วยพื้นฐานย่อยสำหรับมาตรฐานการเข้ารหัส H.264/AVC

การใช้การคืนสภาพมาโครบล็อกภายในและการจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้เพื่อให้วิดิทัศน์ที่บีบอัดมีความต้านทานความผิดพลาด ทั้งสองวิธีนี้จะมีการสร้างบิตที่เพิ่มมากขึ้นกว่าเดิมทำให้บิตที่จัดสรรสำหรับเข้ารหัสเฟรมถัดไปลดลง ซึ่งส่งผลให้วิดิทัศน์ภาษามือที่ได้มีคุณภาพลดลง คุณภาพของวิดิทัศน์ที่ลดลงโดยเฉพาะในบริเวณใบหน้าและมือของผู้ทำภาษามือส่งผลต่อการแปลความหมายที่ถูกต้องภายในวิดิทัศน์ทำได้ยากขึ้น วิทยานิพนธ์นี้้นำการควบคุมอัตราบิตระดับหน่วยพื้นฐานย่อยมาใช้เพื่อปรับปรุงคุณภาพของมาโครบล็อกในบริเวณที่เกี่ยวข้องกับแบบจำลองทางภาษา

อัตราบิต - ความเพี้ยนที่เหมาะสม (rate-distortion optimization) เป็นวิธีที่ใช้ในการควบคุมอัตราบิตระดับหน่วยพื้นฐานย่อย โดยมีการหาค่าควอนไทซ์และโหมดที่เหมาะสมสำหรับการเข้ารหัสมาโครบล็อก ในวิทยานิพนธ์นี้ใช้วิธีอัตราบิต - ความเพี้ยนที่เหมาะสมใน [4] ซึ่งออกแบบมาให้มีความเหมาะสมสำหรับวิดิทัศน์ภาษามือ โดยมีการจัดสรรอัตราบิตให้กับแต่ละมาโครบล็อกไม่เท่ากันโดยพิจารณาจากแบบจำลองทางภาษาที่ให้ความสำคัญกับมาโครบล็อกบริเวณใบหน้ามากที่สุด ส่วนมาโครบล็อกบริเวณมือ ลำตัว และพื้นหลังจะมีความสำคัญลดลงตามลำดับ วิธีอัตราบิต - ความเพี้ยนที่เหมาะสม สามารถคำนวณได้จากสมการที่ (3.16)

$$J = D_{\text{Intell}} + \lambda R \quad (3.16)$$

เมื่อ D_{Intell} คือความเพี้ยนที่มีผลต่อความชัดเจนในการแปลความหมายของวีดิทัศน์

R คือบิตที่ต้องการในการเข้ารหัสของแต่ละมาโครบล็อก

λ คือค่าพารามิเตอร์ลากรองจ์

โดย D_{Intell} ถูกพิจารณาจากความสำคัญตามแบบจำลองทางภาษา สามารถคำนวณได้จากสมการที่ (3.17) เมื่อ k คือบริเวณของใบหน้า มือหรือลำตัว

$$D_{\text{Intell}} = \log_{10}(1 + \alpha_k D_k + D_{\text{NewBG}}) + f(r_f) \quad (3.17)$$

เมื่อ α_k คือ ค่าถ่วงน้ำหนักของหน้า มือหรือลำตัว กำหนดให้มีค่าเท่ากับ 1.6 0.5 และ 0.1 ตามลำดับ

D_k คือ ค่าความเพี้ยนในบริเวณหน้า มือหรือลำตัว

D_{NewBG} คือ ค่าความเพี้ยนจากการเข้ารหัสบริเวณมาโครบล็อกพื้นหลัง

$f(r_f)$ คือ ค่าความเพี้ยนเนื่องจากอัตราเฟรมต่ำ

ค่าความเพี้ยน (D_k) ในบริเวณใบหน้า มือหรือลำตัวเป็นค่าความเพี้ยนเชิงพื้นที่ที่สามารถคำนวณได้ตามสมการที่ (3.18)

$$D_k = \frac{1}{x \cdot y} \sum_{x,y \in MB} (f(x,y) - \tilde{f}(x,y))^2 \quad (3.18)$$

เมื่อ $f(x,y)$ คือ วีดิทัศน์ต้นฉบับที่จุดภาพ (x,y)

$\tilde{f}(x,y)$ คือ วีดิทัศน์ที่สร้างขึ้นที่จุด (x,y)

MB คือ มาโครบล็อกของบริเวณ หน้า มือหรือลำตัว

ความเพี้ยนจากการเข้ารหัสบริเวณมาโครบล็อกพื้นหลัง (D_{NewBG}) เกิดจากการที่ตัวเข้ารหัสเข้ารหัสมาโครบล็อกด้วยโหมดข้าม (skip mode) ในมาโครบล็อกของบริเวณใบหน้าหรือมือในเฟรมก่อนหน้าและเป็นบริเวณของมาโครบล็อกพื้นหลังในเฟรมปัจจุบัน (NewBG) ทำให้บริเวณของหน้าหรือมือมีการตกค้างอยู่ในบริเวณพื้นหลังของเฟรม สามารถคำนวณได้จากสมการที่ (3.19)

$$D_{\text{NewBG}} = \frac{1}{x \cdot y} \sum_{x,y \in MB} (f(x,y) - \tilde{f}(x,y))^2 \quad (3.19)$$

เมื่อ MB เป็นมาโครบล็อกของบริเวณ NewBG

การสูญเสียความเข้าใจง่ายเนื่องจากอัตราเฟรมต่ำ ($f(r_f)$) ที่ระดับคุณภาพเชิงพื้นที่ที่เท่ากัน การลดอัตราเฟรมจะส่งผลต่อคุณภาพของวิดีโอเนื่องจากมีการสูญเสียการเคลื่อนที่ที่สำคัญของท่ามือในวิดีโอ สามารถคำนวณได้จากสมการที่ (3.20)

$$f(r_f) = a_1 * (1 - e^{-a_2 - a_3 r_f}) \quad (3.20)$$

โดยที่ r_f คืออัตราเฟรมในหน่วยเฟรมต่อวินาทีและกำหนดให้ $a_1 = 1.3$, $a_2 = 0.26$ และ $a_3 = 0.34$

โดยค่าพารามิเตอร์ลากรองจ์ (λ) ถูกใช้ในการปรับค่าควอนไทซ์พารามิเตอร์เพื่อให้ได้ค่าใช้จ่าย R-D ที่น้อยที่สุดของแต่ละมาโครบล็อก สามารถคำนวณได้จากสมการที่ (3.21)

$$2 \frac{QP(x)-12}{3} \approx \frac{\lambda}{0.65\alpha(X)} \quad (3.21)$$

โดยที่ $QP(x)$ คือ ควอนไทซ์พารามิเตอร์สำหรับมาโครบล็อก x

$\alpha(x)$ คือ ค่าถ่วงน้ำหนักสำหรับมาโครบล็อก x ในบริเวณใบหน้า มือ ลำตัวและพื้นหลัง กำหนดให้มีค่าเท่ากับ 1.6, 0.5, 0.1 และ 10^{-3} ตามลำดับ

ขั้นตอนวิธีอัตราบิต - ความเพี้ยนที่เหมาะสมจะช่วยให้คุณภาพของวิดีโอดีขึ้นถึงแม้ว่าจะมีการเข้ารหัสที่อัตราบิตต่ำโดยเฉพาะในบริเวณที่เกี่ยวข้องกับแบบจำลองทางภาษาคือบริเวณของใบหน้าจะมีความชัดเจนมากที่สุด บริเวณมือและลำตัวจะมีความชัดเจนรองลงมาส่วนบริเวณพื้นหลังจะมีความชัดเจนต่ำที่สุดเพราะว่าไม่มีผลต่อการเข้าใจความหมายภายในวิดีโอภาษามือของคนพูด

บทที่ 4

ผลการจำลอง

เนื้อหาในบทนี้เป็นผลการจำลองการสื่อสารวีดิทัศน์ไร้สายด้วยการเปลี่ยนค่าพารามิเตอร์ของตัวเข้ารหัส และพารามิเตอร์ของช่องสัญญาณเพื่อศึกษาผลกระทบของความผิดพลาดในช่องสัญญาณที่มีต่อคุณภาพวีดิทัศน์ เมื่อเครื่องมือด้านทานความผิดพลาดและการควบคุมอัตราที่ได้นำเสนอถูกใช้งาน

4.1 พารามิเตอร์ที่ใช้ในการจำลอง

4.1.1 พารามิเตอร์ของช่องสัญญาณไร้สายจำลอง

พารามิเตอร์หลักของช่องสัญญาณไร้สายที่มีการเปลี่ยนแปลงค่าในการจำลองที่มีความผิดพลาดในช่องสัญญาณแตกต่างกันไปคือ ความถี่ดอปเปลอร์ ซึ่งแบ่งการจำลองเป็น 2 กรณีคือ กรณีเฟดดิ้งช้าใช้ความถี่ดอปเปลอร์ 1 เฮิรตซ์ และกรณีเฟดดิ้งเร็วใช้ความถี่ดอปเปลอร์ 40 เฮิรตซ์ โดยความกว้างช่องสัญญาณจะต้องมีความสอดคล้องกับอัตราบิตที่ใช้ในการเข้ารหัส พารามิเตอร์ที่ใช้ในการตั้งค่าซอฟต์แวร์ช่องสัญญาณจำลองเป็นไปตามตารางที่ 4.1

ตารางที่ 4.1 พารามิเตอร์ที่ใช้ในซอฟต์แวร์ช่องสัญญาณจำลอง

พารามิเตอร์	ประเภท/ค่าที่ใช้
Multiple Access	TDMA
Modulation	QPSK
Channel rate	64, 96, 128, 160 kbps
Maximum Doppler Frequency	1Hz , 40Hz
Transmitted Signal Power	15 dB
Time delay spread	¼ of symbol period
Power delay profile	2-ray with equal power
Antenna Diversity	1

4.1.2 พารามิเตอร์ของตัวเข้ารหัส/ตัวถอดรหัสวีดิทัศน์ H.264

พารามิเตอร์ของตัวเข้ารหัส/ตัวถอดรหัสแบ่งออกเป็น 2 ส่วนได้แก่ ส่วนตัวเข้ารหัส และ ส่วนตัวถอดรหัส มีรายละเอียดดังนี้

4.1.2.1 ตัวเข้ารหัส

ตัวเข้ารหัสที่ใช้เป็นตัวเข้ารหัส H.264 ของ JVT รุ่น JM18.4 [25] เนื่องจากไม่มีมาตรฐานของวีดิทัศน์ภาษามือที่ใช้สำหรับทดสอบ ในวิทยานิพนธ์นี้จึงใช้วีดิทัศน์ภาษามือจำนวน 4 ลำดับภาพ ได้แก่ Sign I, Sign II, Sign III และ Sign IV จาก [7] โดยจะถูกแยกบริเวณของ ใบหน้า มือ ลำตัวและพื้นหลังก่อนทำการเข้ารหัสโดยใช้วิธีแบบจำลองสีผิว [34]

ข้อกำหนดหลักในการเข้ารหัสมีดังต่อไปนี้

1. วีดิทัศน์ภาษามือที่อัตราบิตต่ำสำหรับโทรศัพท์วีดิทัศน์ใช้ความละเอียดที่ขนาด CIF (352x288 จุดภาพ) เนื่องจากให้รายละเอียดของมือที่ชัดในระบบสี $YCbCr$ แบบ 4:2:0
2. จำนวนเฟรมที่เข้ารหัสของแต่ละลำดับภาพ 100 เฟรม เข้ารหัสที่อัตราเฟรม 15 เฟรมต่อวินาที
3. ภาพเป็นชนิด I และ P เท่านั้น ไม่มีการใช้ภาพชนิด B
4. จำนวนกลุ่มสไลซ์ต่อเฟรมกำหนดค่าอยู่ที่ 8 กลุ่มสไลซ์
5. ไม่มีการกระโดดของเฟรมเนื่องจากการทำงานของตัวควบคุมอัตราในตัวเข้ารหัส
6. การควบคุมอัตราบิตถูกใช้งานที่ 64, 96, 128, 160 Kbps
7. โพรโตคอลที่ใช้ในการเข้ารหัสเป็นแบบ RTP

ที่ตัวเข้ารหัสมีการปรับให้สามารถเลือกโหมดการเข้ารหัสของมาโครบล็อกเฉพาะโหมดการเข้ารหัสภายในได้ อย่างไรก็ตามตัวเข้ารหัสรุ่นนี้ไม่รองรับการเข้ารหัสโดยใช้แผนที่ยกกลุ่มสไลซ์แบบชัดแจ้งที่เปลี่ยนไปในทุกเฟรม จึงมีการปรับปรุงให้ตัวเข้ารหัสสามารถปรับพารามิเตอร์ภาพให้เป็นปัจจุบันเพื่อรองรับแผนที่ยกกลุ่มสไลซ์แบบชัดแจ้งสำหรับแต่ละเฟรม

4.1.2.2 ตัวถอดรหัส

ตัวถอดรหัสที่ใช้เป็นรุ่นเดียวกันกับตัวเข้ารหัส ตัวถอดรหัสถูกแก้ไขเพื่อจำลองผลกระทบของความผิดพลาดที่มีต่อวีดิทัศน์ เมื่อตัวถอดรหัสพบว่าข้อมูลมาโครบล็อกที่กำลังถอดรหัสผิดพลาด ก็จะไม่สามารถถอดรหัสมาโครบล็อกนั้นและมาโครบล็อกต่อจากนั้นภายในกลุ่มสไลซ์เดียวกัน ตัวถอดรหัสเพิ่มการปกปิดความผิดพลาดแบบไม่มีการชดเชยการเคลื่อนที่สำหรับมาโครบล็อกที่ถอดรหัสไม่ได้คือ

ให้ใช้ข้อมูลของมาโครบล็อกจากเฟรมก่อนหน้าที่ตำแหน่งเดียวกันแสดงผลแทนในวีดิทัศน์ที่ถอดรหัสได้

4.2 การวัดประสิทธิภาพของเทคนิคที่ได้นำเสนอ

ค่าที่ใช้วัดเพื่อเปรียบเทียบประสิทธิภาพของแต่ละวิธีหรือเทคนิคที่ใช้ในการจำลองนี้ใช้ค่าเชิงประวิสัยในการวัดมี 2 แบบได้แก่ค่าคุณภาพวีดิทัศน์ PSNR และค่าความชัดเจนในการแปลความหมายของวีดิทัศน์ภาษามือ CIM

4.2.1 ค่าคุณภาพวีดิทัศน์ PSNR

ค่าคุณภาพวีดิทัศน์ PSNR นอกจะขึ้นกับความผิดพลาดที่เกิดขึ้นในกระแสบิตแล้ว ยังมีปัจจัยในเรื่องของกระบวนการภายในของตัวตัวเข้ารหัส/ตัวถอดรหัส ทั้งทางด้านเข้ารหัสเช่นการควบคุมอัตรา การทำนายการเคลื่อนที่ และทางด้านตัวถอดรหัสเช่น การปกปิดความผิดพลาด การชดเชยการเคลื่อนที่

4.2.2 ค่าความชัดเจนในการแปลความหมายของวีดิทัศน์ภาษามือ CIM

ค่าความชัดเจนในการแปลความหมายเป็นค่าที่บ่งบอกถึงค่าความเพี้ยนในบริเวณที่ใช้สำหรับการแปลความหมายของภาษามือ ซึ่งค่าความชัดเจนในการแปลความหมายยิ่งสูงจะทำให้มีโอกาสที่จะสามารถแปลความหมายได้ถูกต้องด้วยเช่นกัน รายละเอียดการคำนวณกล่าวไว้ในหัวข้อที่

4.3 ผลการจำลอง

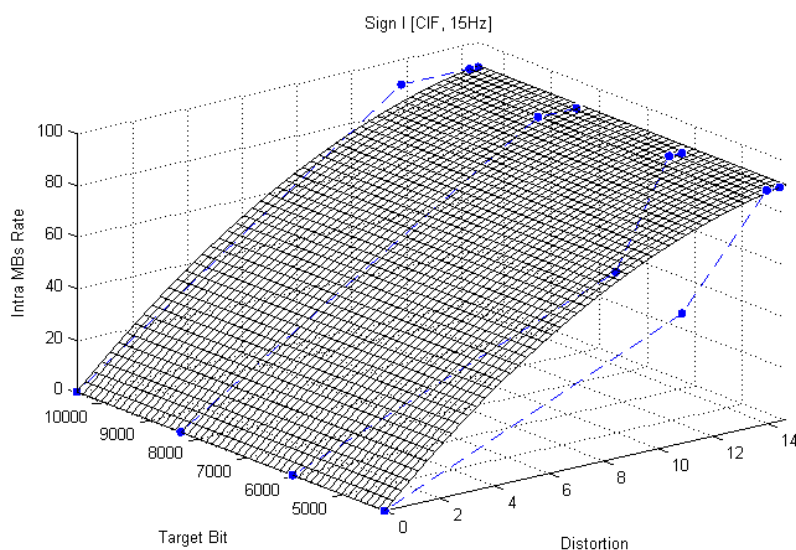
4.3.1 การหาอัตราการคืนสภาพมาโครบล็อกภายในที่เหมาะสม

การทำการทดลองอย่างละเอียดถูกนำไปใช้ในการหาอัตราการคืนสภาพมาโครบล็อกภายในที่เหมาะสมสำหรับวีดิทัศน์ภาษามือ อัตราการคืนสภาพมาโครบล็อกภายในที่เหมาะสมคืออัตราร้อยละของจำนวนมาโครบล็อกภายในที่ใช้ในการเข้ารหัสบริเวณใบหน้าและมือต่อจำนวนมาโครบล็อกทั้งหมดของใบหน้าและมือในหนึ่งเฟรม โดยทำการทดสอบกับอัตราร้อยละของจำนวนมาโครบล็อกภายในที่แตกต่างกันออกไปแล้วเลือกเพียงหนึ่งเดียวที่ให้ค่าเฉลี่ย PSNR สูงที่สุดสำหรับแต่ละอัตราบิตและความผิดพลาดที่แตกต่างกันออกไป ข้อกำหนดที่ใช้ในการทดสอบสรุปตารางที่ 4.2

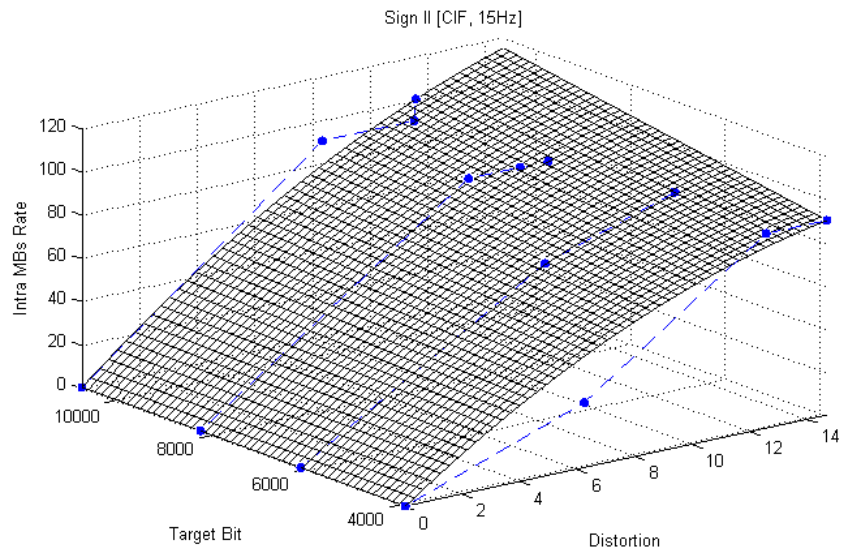
ตารางที่ 4.2 ข้อกำหนดในการทดสอบหาอัตราคืนสภาพภายในที่เหมาะสม

เงื่อนไข	ค่า
ลำดับภาพ	Sign I, Sign II, Sign III และ Sign IV
ขนาดภาพ	CIF
อัตราเฟรม (เฟรมต่อวินาที)	15
อัตราบิต (กิโลบิตต่อวินาที)	64, 96, 128, 160
จำนวนมาโครบล็อกภายใน (อัตราร้อยละ)	10%, 20%, 30%, ...,100%
ความผิดพลาดที่ความถี่ต่อเพลอร์ (เฮิร์ตซ์)	1, 10, 40

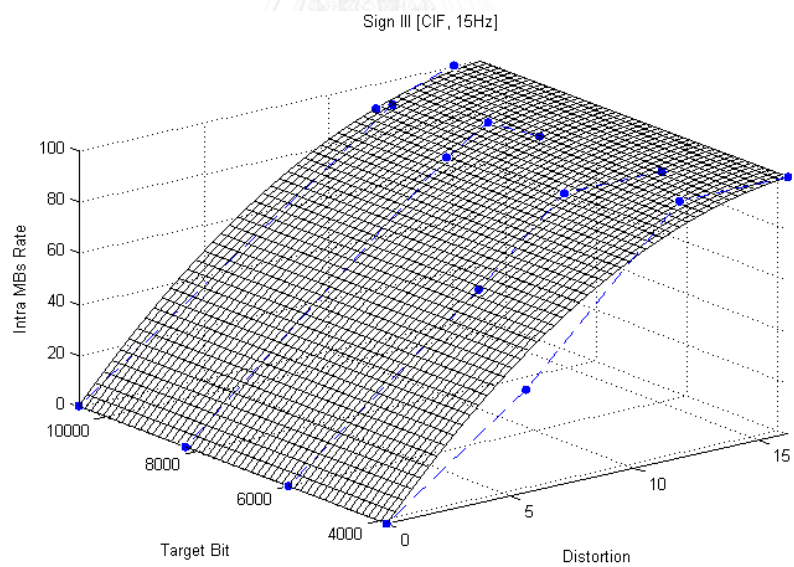
ใน 1 ลำดับภาพจะได้ค่าตัวอย่างทั้งหมด 12 ค่า จากการพิจารณาที่ 4 อัตราบิตและ 3 ความถี่ต่อเพลอร์ของช่องสัญญาณ ค่าตัวอย่างจะแสดงความสัมพันธ์ระหว่างอัตราร้อยละของจำนวนมาโครบล็อกภายในที่ใช้ในการเข้ารหัสที่ให้ค่า PSNR สูงที่สุดต่อค่าความเพี้ยนในบริเวณใบหน้าและมือกับบิตเป้าหมายเฉลี่ยสำหรับทุกเฟรมที่ใช้ในการเข้ารหัสลำดับภาพ รูปที่ 4.1 ถึงรูปที่ 4.4 แสดงค่าตัวอย่างที่ได้จากการทดลองอย่างละเอียดและการปรับเส้นโค้ง (Curve fitting) ของแต่ละลำดับภาพ



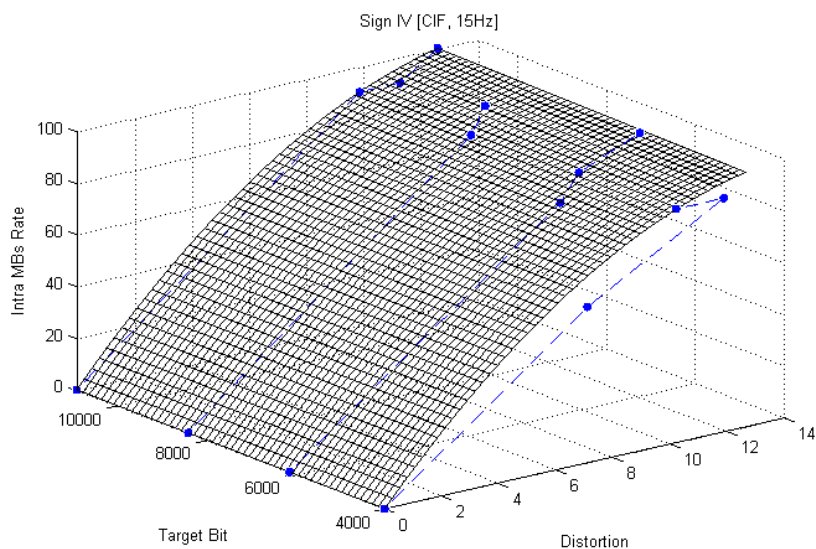
รูปที่ 4.1 ค่าตัวอย่างที่ได้จากการทดลองอย่างละเอียดและการปรับเส้นโค้งของลำดับภาพ Sign I



รูปที่ 4.2 ค่าตัวอย่างที่ได้จากการทดลองอย่างละเอียดและการปรับเส้นโค้งของลำดับภาพ Sign II



รูปที่ 4.3 ค่าตัวอย่างที่ได้จากการทดลองอย่างละเอียดและการปรับเส้นโค้งของลำดับภาพ Sign III



รูปที่ 4.4 ค่าตัวอย่างที่ได้จากการทดลองอย่างละเอียดและการปรับเส้นโค้งของลำดับภาพ Sign IV

4.3.2 การจัดสรรการคืนสภาพมาโครบล็อกภายในบริเวณใบหน้าและมือของผู้ทำท่าภาษามือ

ผลการทดลองในส่วนนี้เป็นผลจากการใช้มาโครบล็อกภายในในการเข้ารหัสบริเวณใบหน้าและมือของผู้ทำท่าภาษามือในวิดีโอเพื่อเป็นการลดผลของความผิดพลาดจากช่องสัญญาณที่มีต่อคุณภาพของวิดีโอ โดยจำนวนของมาโครบล็อกภายในที่เหมาะสมสำหรับการเข้ารหัสวิดีโอภาษามือหาได้จากแบบจำลองอัตราคืนสภาพมาโครบล็อกภายใน ตารางที่ 4.3 และ ตารางที่ 4.4 เปรียบเทียบค่าเฉลี่ย PSNR บริเวณใบหน้าและมือของผู้ทำท่าภาษามือระหว่างการใช้มาโครบล็อกภายในตามแบบจำลองอัตราคืนสภาพมาโครบล็อกภายใน (Intra_Model) ตามหัวข้อที่ 3.1.2 กับวิธี AAIR ที่ใช้โหมดการเข้ารหัสภายในเข้ารหัสทุกมาโครบล็อกของใบหน้าและมือทุกๆ 3 เฟรม เมื่อทดสอบภายใต้การจำลองความผิดพลาดจากช่องสัญญาณกรณีเฟตติงช้าและกรณีเฟตติงเร็ว

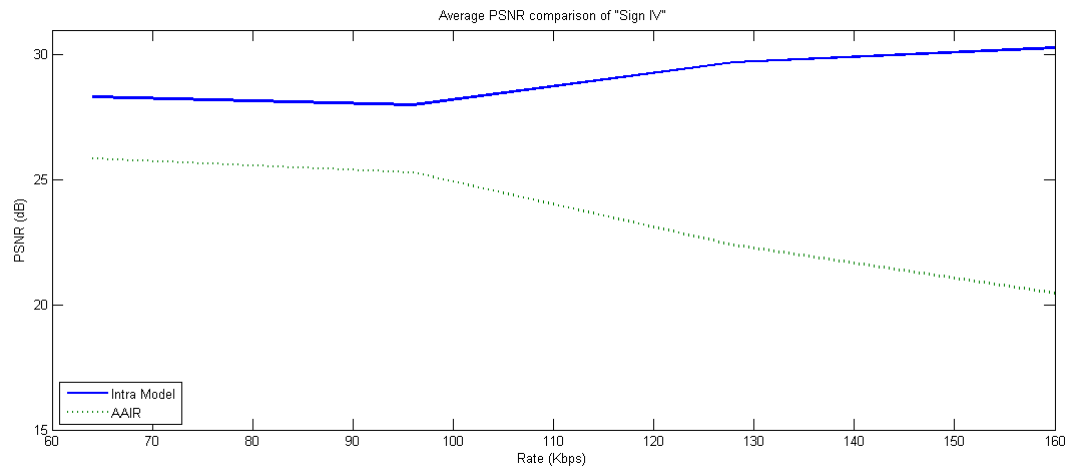
ตารางที่ 4.3 เปรียบเทียบค่าเฉลี่ย PSNR บริเวณใบหน้าและมือของผู้ทำท่าภาษามือกรณีเฟดดิ้งช้า

เฟดดิ้งช้า								
วิธี \ ลำดับภาพ	64 kbps				96 kbps			
	Sign I	Sign II	Sign III	Sign IV	Sign I	Sign II	Sign III	Sign IV
Intra_Model	27.00	28.50	29.10	28.32	28.38	28.76	28.23	28.00
AAIR	27.70	27.88	23.93	25.86	25.16	31.04	25.45	25.29
วิธี \ ลำดับภาพ	128 kbps				160 kbps			
	Sign I	Sign II	Sign III	Sign IV	Sign I	Sign II	Sign III	Sign IV
Intra_Model	27.15	28.36	25.94	29.70	28.73	30.25	26.54	30.28
AAIR	23.10	24.03	20.33	22.38	21.69	23.19	19.88	20.46

ตารางที่ 4.4 เปรียบเทียบค่าเฉลี่ย PSNR บริเวณใบหน้าและมือของผู้ทำท่าภาษามือกรณีเฟดดิ้งเร็ว

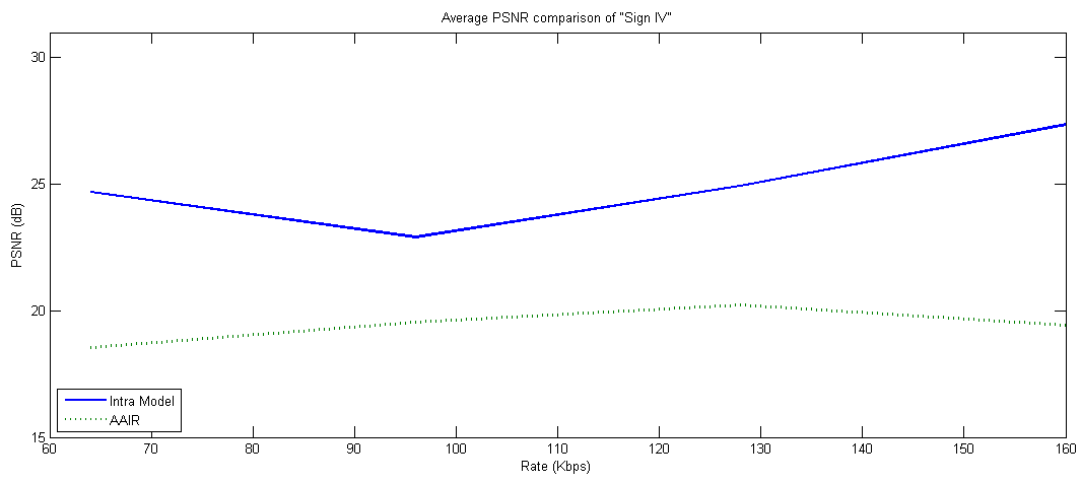
เฟดดิ้งเร็ว								
วิธี \ ลำดับภาพ	64 kbps				96 kbps			
	Sign I	Sign II	Sign III	Sign IV	Sign I	Sign II	Sign III	Sign IV
Intra_Model	22.44	19.87	18.40	22.91	21.59	20.73	18.87	24.70
AAIR	19.92	15.94	18.26	18.54	20.79	16.69	18.12	19.54
วิธี \ ลำดับภาพ	128 kbps				160 kbps			
	Sign I	Sign II	Sign III	Sign IV	Sign I	Sign II	Sign III	Sign IV
Intra_Model	20.64	20.34	17.37	24.93	23.25	20.26	17.99	27.37
AAIR	19.78	15.91	18.39	20.22	18.05	16.43	20.18	19.42

ผลการทดลองแสดงให้เห็นว่าการใช้มาโครบล็อกภายในด้วยวิธีที่นำเสนอช่วยให้คุณภาพของวิดีโอที่เก็บทุกกรณีดีขึ้น ที่เฟดดิ้งช้าเมื่อทำการเข้ารหัสที่อัตราบิต 160 kbps คุณภาพวิดีโอของลำดับภาพ Sign IV สูงขึ้น 9.82 dB และเมื่อทำการเปรียบเทียบที่เฟดดิ้งเร็วกว่าที่อัตราบิต 160 kbps คุณภาพวิดีโอของลำดับภาพ Sign IV สูงขึ้น 7.95 dB เมื่อเทียบกับวิธี AAIR การวัดโดยคุณภาพของวิดีโอโดยใช้ค่า PSNR จะขึ้นอยู่กับตำแหน่งของความผิดพลาดในช่องสัญญาณและเทคนิคการปกปิดความผิดพลาดที่ใช้ รูปที่ 4.5 และรูปที่ 4.6 กราฟค่าเฉลี่ย PSNR ของลำดับภาพ Sign IV กรณีเฟดดิ้งช้าและเร็วตามลำดับ รูปที่ 4.7 และรูปที่ 4.8 แสดงคุณภาพของวิดีโอในลำดับภาพ Sign IV ที่ถอดรหัส



รูปที่ 4.5 กราฟค่าเฉลี่ย PSNR เปรียบเทียบระหว่างวิธี Intra_Model กับ AAIR ของลำดับภาพ

Sign IV กรณีเฟดดิ้งช้า



รูปที่ 4.6 กราฟค่าเฉลี่ย PSNR เปรียบเทียบระหว่างวิธี Intra_Model กับ AAIR ของลำดับภาพ

Sign IV กรณีเฟดดิ้งเร็ว



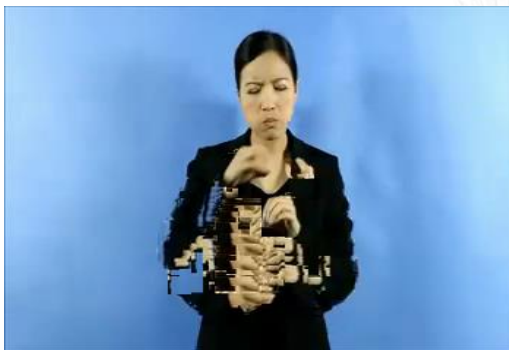
(ก)



(ข)

รูปที่ 4.7 เฟรมที่ 57 ของลำดับภาพ Sign IV ที่อัตราบิต 160 kbps กรณีเฟดดิ้งช้าโดย

(ก) Intra_Model (ข) AAIR



(ก)



(ข)

รูปที่ 4.8 เฟรมที่ 34 ของลำดับภาพ Sign IV ที่อัตราบิต 160 kbps กรณีเฟดดิ้งเร็วโดย

(ก) Intra_Model (ข) AAIR

4.3.3 การจัดสรรการคืนสภาพมาโครบล็อกภายในบริเวณลำตัวของผู้ทำท่าภาษามือ

ผลการทดลองในส่วนนี้เป็นผลของการใช้แผนการคืนสภาพมาโครบล็อกภายใน (Intra_Scheme) กล่าวคือมีการใช้การคืนสภาพมาโครบล็อกภายในบริเวณใบหน้า มือและลำตัวของ ผู้ทำท่าภาษามือตามหัวข้อที่ 3.1.2 และ 3.1.3 เพื่อเป็นการลดผลของความผิดพลาดจากช่องสัญญาณ ผลการทดลองแสดงคุณภาพของวิดีโอที่คืนสภาพภาษามือหลังจากมีการปรับปรุงคุณภาพในบริเวณลำตัว เพิ่มเติมจากการทดลองในหัวข้อที่ 4.3.2

ตารางที่ 4.5 เปรียบเทียบค่าเฉลี่ย PSNR บริเวณใบหน้า มือและลำตัวกรณีเฟตดิ่งช้า

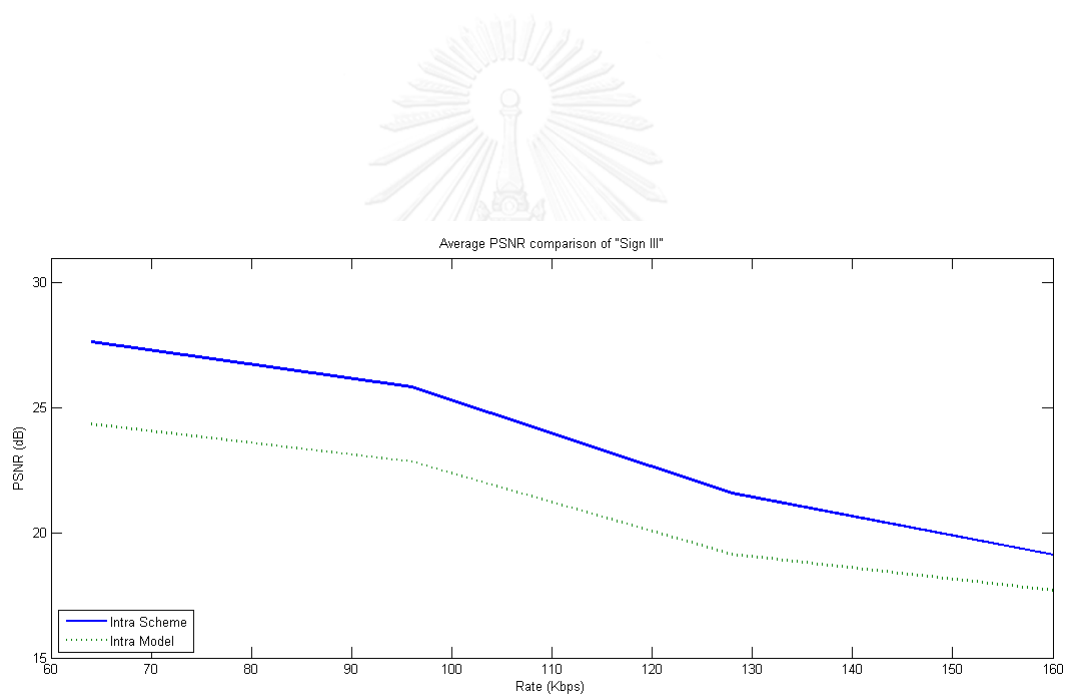
เฟตดิ่งช้า								
วิธี \ ลำดับภาพ	64 kbps				96 kbps			
	Sign I	Sign II	Sign III	Sign IV	Sign I	Sign II	Sign III	Sign IV
Intra_Scheme	24.24	22.48	27.65	21.40	24.15	26.53	25.84	23.69
Intra_Model	22.13	22.88	24.35	20.59	22.08	26.01	22.85	23.24
วิธี \ ลำดับภาพ	128 kbps				160 kbps			
	Sign I	Sign II	Sign III	Sign IV	Sign I	Sign II	Sign III	Sign IV
Intra_Scheme	18.73	25.55	21.59	24.49	20.10	26.81	19.11	24.45
Intra_Model	17.58	23.51	19.14	22.94	19.36	24.09	17.70	23.35

ตารางที่ 4.6 เปรียบเทียบค่าเฉลี่ย PSNR บริเวณใบหน้า มือและลำตัวกรณีเฟตดิ่งเร็ว

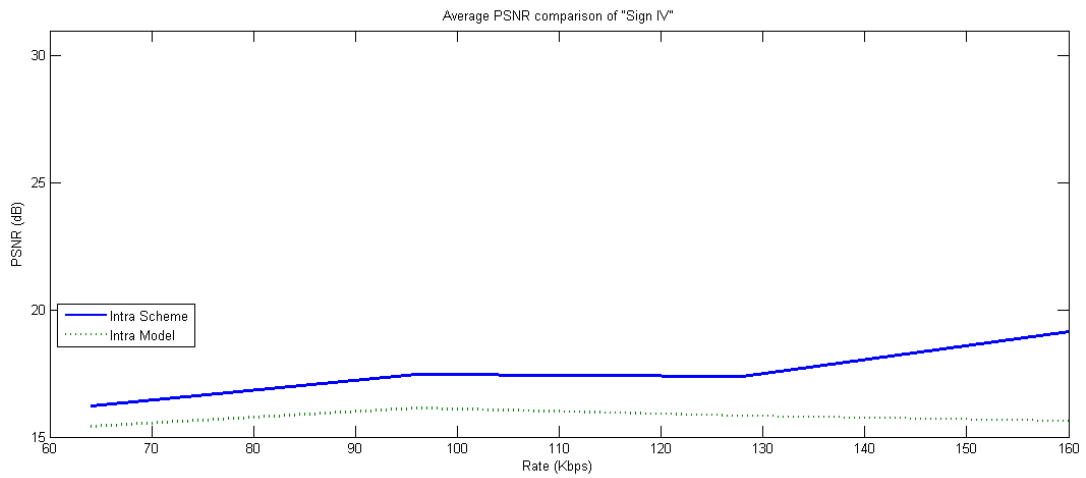
เฟตดิ่งเร็ว								
วิธี \ ลำดับภาพ	64 kbps				96 kbps			
	Sign I	Sign II	Sign III	Sign IV	Sign I	Sign II	Sign III	Sign IV
Intra_Scheme	15.82	15.23	15.88	16.22	14.67	16.76	16.33	17.46
Intra_Model	14.49	15.30	14.80	15.42	14.27	16.62	15.33	16.14
วิธี \ ลำดับภาพ	128 kbps				160 kbps			
	Sign I	Sign II	Sign III	Sign IV	Sign I	Sign II	Sign III	Sign IV
Intra_Scheme	14.39	16.92	15.79	17.38	14.51	17.32	14.98	19.15
Intra_Model	13.46	16.34	15.08	15.84	13.69	17.16	14.47	15.64

ตารางที่ 4.5 และ ตารางที่ 4.6 เปรียบเทียบค่าเฉลี่ย PSNR บริเวณใบหน้า มือและลำตัวใน วิดีทัศน์เมื่อมีการปรับปรุงคุณภาพบริเวณลำตัวโดยเพิ่มจำนวนมาโครบล็อกภายในจากวิธีก่อนหน้านี้ เทียบกับการการใช้มาโครบล็อกภายในตามแบบจำลองอัตราคินสภาพมาโครบล็อกภายใน (Intra_Model) ภายใต้การจำลองความผิดพลาดจากช่องสัญญาณกรณีเฟตดิ่งช้าและกรณีเฟตดิ่งเร็ว

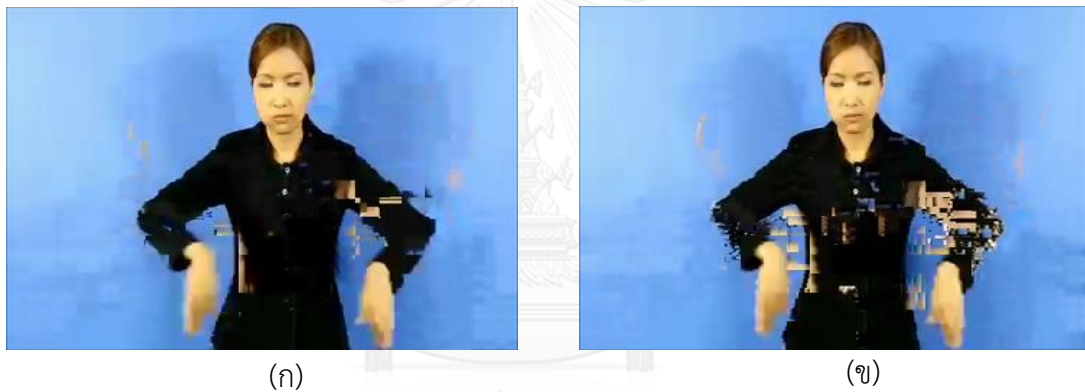
ผลการทดลองแสดงให้เห็นว่าแผนการคืนสภาพมาโครบล็อกภายใน (Intra_Scheme) ช่วย
ให้คุณภาพของวิดีโอในบริเวณใบหน้า มือและลำตัวในทุกกรณีเพิ่มขึ้น ที่เฟรมดิ่งช้าเมื่อทำการ
เข้ารหัสที่อัตราบิต 96 kbps คุณภาพวิดีโอของลำดับภาพ Sign III สูงขึ้น 2.99 dB และเมื่อทำการ
เปรียบเทียบที่เฟรมดิ่งเร็วพบว่าที่อัตราบิต 160 kbps คุณภาพวิดีโอของลำดับภาพ Sign IV สูงขึ้น
3.51 dB เมื่อเทียบกับวิธีการเข้ารหัสโดยใช้มาโครบล็อกภายในตามแบบจำลองอัตราคืนสภาพมาโคร
บล็อกภายใน รูปที่ 4.9 และ รูปที่ 4.10 กราฟค่าเฉลี่ย PSNR ของลำดับภาพ Sign III กรณีเฟรมดิ่งช้า
และลำดับภาพ Sign IV กรณีเฟรมดิ่งเร็วตามลำดับ รูปที่ 4.11 ถึง รูปที่ 4.14 แสดงคุณภาพของวิดีโอ
ทัศน์ในลำดับภาพ Sign III และ Sign IV ตามลำดับ



รูปที่ 4.9 กราฟค่าเฉลี่ย PSNR เปรียบเทียบระหว่างวิธี Intra_Scheme กับ Intra_Model ของลำดับ
ภาพ Sign III กรณีเฟรมดิ่งช้า



รูปที่ 4.10 กราฟค่าเฉลี่ย PSNR เปรียบเทียบระหว่างวิธี Intra_Scheme กับ Intra_Model ของลำดับภาพ Sign III กรณีเฟดดิ้งเร็ว



รูปที่ 4.11 เฟรมที่ 80 ของลำดับภาพ Sign III ที่อัตราบิต 96 kbps กรณีเฟดดิ้งช้าโดย (ก) Intra_Scheme (ข) Intra_Model



รูปที่ 4.12 เฟรมที่ 85 ของลำดับภาพ Sign III ที่อัตราบิต 96 kbps กรณีเฟดดิ้งเร็วโดย (ก) Intra_Scheme (ข) Intra_Model



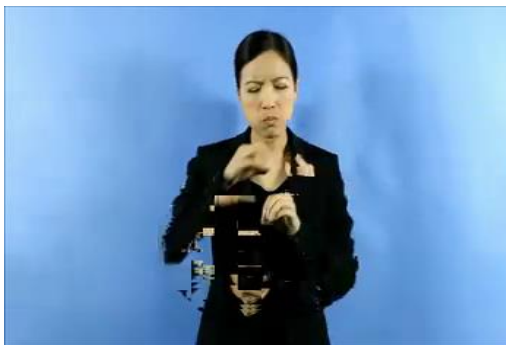
(ก)



(ข)

รูปที่ 4.13 เฟรมที่ 67 ของลำดับภาพ Sign IV ที่อัตราบิต 160 kbps กรณีเฟดดิ้งช้าโดย

(ก) Intra_Scheme (ข) Intra_Model



(ก)



(ข)

รูปที่ 4.14 เฟรมที่ 34 ของลำดับภาพ Sign IV ที่อัตราบิต 160 kbps กรณีเฟดดิ้งเร็วโดย

(ก) Intra_Scheme (ข) Intra_Model

4.3.4 การใช้การจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้

ในหัวข้อที่ 3.2 มีการใช้การจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้เพื่อลดผลของความผิดพลาดที่เกิดขึ้นจากการแพร่กระจายภายในเฟรม ในวิทยานิพนธ์นี้ได้นำเสนอการปรับปรุงกระบวนการจัดสรรมาโครบล็อกเข้าสู่กลุ่มสไลซ์เพื่อให้ความยืดหยุ่นจากการปกปิดความผิดพลาดของบริเวณใบหน้ามือและลำตัวเท่ากันหรือใกล้เคียงกันมากที่สุดในแต่ละกลุ่มสไลซ์ เพื่อเป็นการลดความน่าจะเป็นที่ความเพี้ยนจะไปสะสมอยู่ในกลุ่มสไลซ์ใดกลุ่มสไลซ์หนึ่งมากเกินไปซึ่งจะส่งผลกระทบต่อคุณภาพวีดิทัศน์สูง เพื่อเป็นการเปรียบเทียบประสิทธิภาพของการปรับปรุงการจัดสรรมาโครบล็อกเข้าสู่กลุ่มสไลซ์ ค่าเบี่ยงเบนมาตรฐาน (Standard Deviation) ถูกนำมาใช้ในการวัดการกระจายตัวของค่าความเพี้ยนในแต่ละกลุ่มสไลซ์ตารางที่ 4.7 และตารางที่ 4.8 แสดงค่าเบี่ยงเบนมาตรฐานระหว่างเทคนิคที่ได้

นำเสนอ (Improved_sort) กับการจัดสรรมาโครบล็อกเข้าสู่กลุ่มสไลซ์โดยพิจารณาค่าความเพี้ยนแบบเดิม (Distortion_sort) ที่กรณีเฟตติงช้าและกรณีเฟตติงเร็วตามลำดับ

ตารางที่ 4.7 เปรียบเทียบค่าเบี่ยงเบนมาตรฐานของค่าความเพี้ยนที่กรณีเฟตติงช้า

เฟตติงช้า								
ลำดับภาพ วิธี	64 kbps				96 kbps			
	Sign I	Sign II	Sign III	Sign IV	Sign I	Sign II	Sign III	Sign IV
Improved_sort	0.50	6.64	1.01	0.25	0.26	0.65	0.58	4.09
Distortion_sort	24.89	21.00	15.77	12.27	20.33	15.16	18.12	16.91
ลำดับภาพ วิธี	128 kbps				160 kbps			
	Sign I	Sign II	Sign III	Sign IV	Sign I	Sign II	Sign III	Sign IV
Improved_sort	0.33	0.58	1.58	7.39	0.45	1.33	4.18	5.99
Distortion_sort	44.23	20.07	30.85	18.38	30.31	20.17	36.90	15.48

ตารางที่ 4.8 เปรียบเทียบค่าเบี่ยงเบนมาตรฐานของค่าความเพี้ยนที่กรณีเฟตติงเร็ว

เฟตติงเร็ว								
ลำดับภาพ วิธี	64 kbps				96 kbps			
	Sign I	Sign II	Sign III	Sign IV	Sign I	Sign II	Sign III	Sign IV
Improved_sort	0.49	0.44	0.41	0.35	0.40	0.32	0.47	0.35
Distortion_sort	48.06	55.50	57.22	28.25	56.08	52.42	51.42	21.96
ลำดับภาพ วิธี	128 kbps				160 kbps			
	Sign I	Sign II	Sign III	Sign IV	Sign I	Sign II	Sign III	Sign IV
Improved_sort	0.42	0.44	0.42	0.33	0.42	0.38	0.44	0.48
Distortion_sort	58.41	54.66	59.98	23.40	57.68	50.53	59.81	20.92

ตารางที่ 4.9 เปรียบเทียบค่าเฉลี่ย CIM ในบริเวณของผู้ทำท่าภาษามือกรณีเฟตดิ่งช้า

เฟตดิ่งช้า								
วิธี	64 kbps				96 kbps			
	Sign I	Sign II	Sign III	Sign IV	Sign I	Sign II	Sign III	Sign IV
FMO_Imp+ Intra_Scheme	2.33	2.33	2.42	2.35	2.33	2.41	2.41	2.40
FMO_Dis+ Intra_Scheme	2.35	2.44	2.37	2.38	2.33	2.47	2.41	2.41
Intra_Scheme	2.14	2.13	2.25	2.06	2.15	2.28	2.22	2.07
วิธี	128 kbps				160 kbps			
	Sign I	Sign II	Sign III	Sign IV	Sign I	Sign II	Sign III	Sign IV
FMO_Imp+ Intra_Scheme	2.11	2.29	2.07	2.23	2.13	2.34	1.90	2.26
FMO_Dis+ Intra_Scheme	2.19	2.26	2.05	2.17	2.16	2.29	1.96	2.19
Intra_Scheme	1.78	2.10	1.84	2.10	1.86	2.16	1.70	2.08

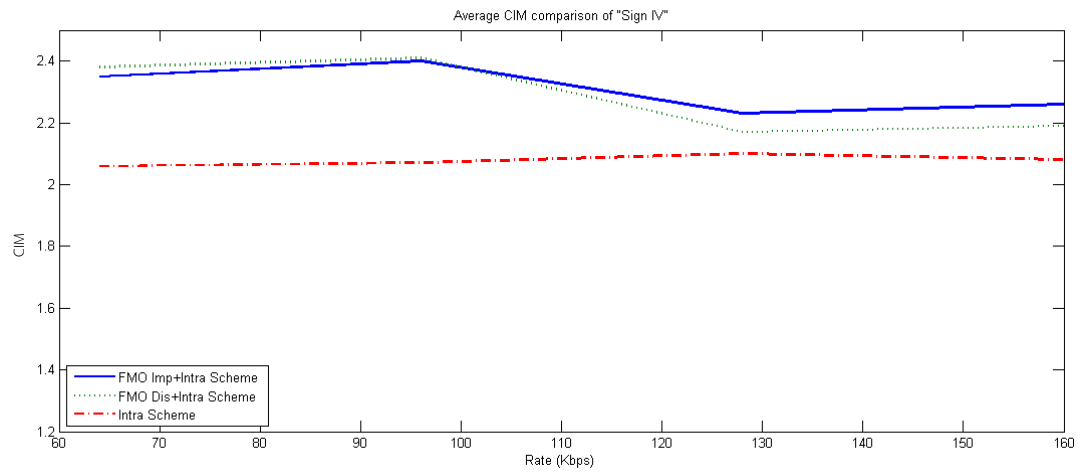
จากผลการทดลองพบว่าค่าเบี่ยงเบนมาตรฐานของวิธีที่นำเสนอมีค่าน้อยกว่าในทุกกรณีแสดงว่าค่าความเพี้ยนรวมของบริเวณหน้า มือและลำตัวในแต่ละกลุ่มสไลซ์มีค่าใกล้เคียงกันมากกว่าวิธีเดิม

วัตถุประสงค์ของการปรับปรุงคุณภาพของวีดิทัศน์ภาษามือคือเพื่อให้มีความชัดเจนในการแปลความหมายได้ถูกต้อง การทดลองตั้งแต่หัวข้อนี้เป็นต้นไปเป็นการปรับปรุงคุณภาพวีดิทัศน์ทั้งในส่วนบริเวณใบหน้า มือและลำตัวซึ่งสอดคล้องกับการวัดความชัดเจนในการแปลความหมายของวีดิทัศน์ภาษามือ จึงใช้การวัดความชัดเจนในการแปลความหมายของวีดิทัศน์เพื่อทดสอบประสิทธิภาพของเทคนิคที่ได้นำเสนอ ตารางที่ 4.9 และ ตารางที่ 4.10 เปรียบเทียบค่าเฉลี่ย CIM ของวีดิทัศน์เมื่อมีการใช้การจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้โดยใช้การจัดสรรมาโครบล็อกเข้าสู่กลุ่มสไลซ์ทั้งแบบปรับปรุงการจัดสรรมาโครบล็อกเข้าสู่กลุ่มสไลซ์ (FMO_Imp) และแบบเดิม (FMO_Dis) ร่วมกับแผนการคืนสภาพมาโครบล็อกภายใน (Intra_Scheme) เทียบกับการเข้ารหัสโดยใช้แผนการคืนสภาพมาโครบล็อกภายใน (Intra_Scheme) ภายใต้การจำลองความผิดพลาดจากช่องสัญญาณกรณีเฟตดิ่งช้าและกรณีเฟตดิ่งเร็ว

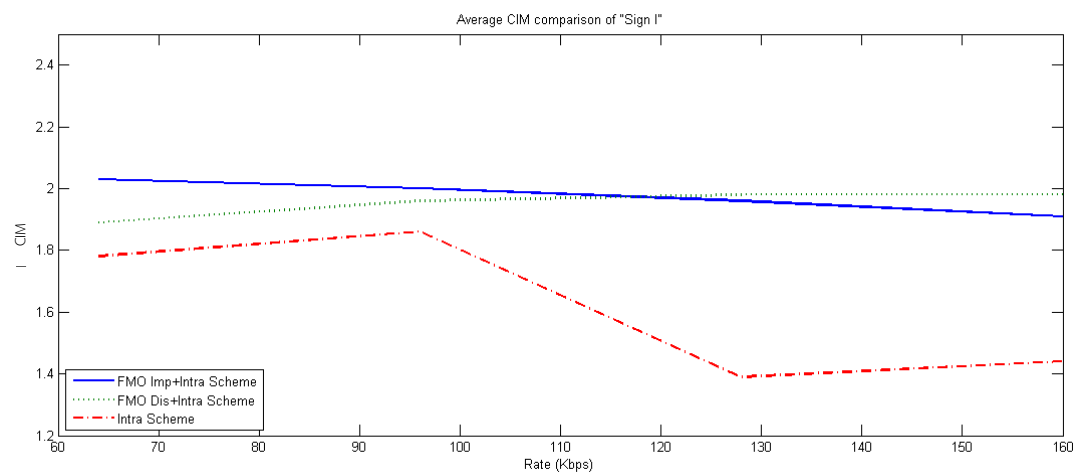
ตารางที่ 4.10 เปรียบเทียบค่าเฉลี่ย CIM ในบริเวณของผู้ทำท่าภาษามือกรณีเฟตติงเร็ว

เฟตติงเร็ว								
ลำดับภาพ วิธี	64 kbps				96 kbps			
	Sign I	Sign II	Sign III	Sign IV	Sign I	Sign II	Sign III	Sign IV
FMO_Imp+ Intra_Scheme	2.03	1.96	1.86	2.11	2.00	1.78	1.77	2.13
FMO_Dis+ Intra_Scheme	1.89	1.87	1.86	2.07	1.96	1.86	1.77	2.01
Intra_Scheme	1.78	2.10	1.84	2.10	1.86	2.16	1.70	2.08
ลำดับภาพ วิธี	128 kbps				160 kbps			
	Sign I	Sign II	Sign III	Sign IV	Sign I	Sign II	Sign III	Sign IV
FMO_Imp+ Intra_Scheme	1.96	1.76	1.77	2.14	1.91	1.68	1.77	2.22
FMO_Dis+ Intra_Scheme	1.98	1.75	1.76	2.01	1.98	1.72	1.70	2.29
Intra_Scheme	1.39	1.51	1.32	1.71	1.44	1.55	1.34	1.88

จากผลการทดลองเมื่อทำการเปรียบเทียบระหว่างวิธีการจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้ โดยปรับปรุงการจัดสรรมาโครบล็อกร่วมกับแผนการคืนสภาพมาโครบล็อกภายใน (FMO_Imp+ Intra_Scheme) เทียบกับวิธีการจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้โดยใช้การจัดสรรมาโครบล็อกแบบเดิมร่วมกับแผนการคืนสภาพมาโครบล็อกภายใน (FMO_Dis+Intra_Scheme) พบว่าการเข้ารหัสที่อัตราบิต 160 kbps ความชัดเจนในการแปลความหมายของวิดีโอที่ลำดับภาพ Sign IV สูงขึ้น 0.07 และที่อัตราบิต 64 kbps ความชัดเจนในการแปลความหมายของวิดีโอที่ลำดับภาพ Sign I สูงขึ้น 0.14 ทั้งในกรณีเฟตติงช้าและเฟตติงเร็วตามลำดับ เมื่อเทียบกับวิธีแผนการคืนสภาพมาโครบล็อกภายใน (Intra_Scheme) พบว่าการเข้ารหัสที่อัตราบิต 96 kbps ของวิดีโอที่ลำดับภาพ Sign IV และที่อัตราบิต 128 kbps ของวิดีโอที่ลำดับภาพ Sign I ความชัดเจนในการแปลความหมายสูงขึ้น 0.33 ในกรณีเฟตติงช้า สำหรับกรณีเฟตติงเร็วอัตราบิต 128 kbps ความชัดเจนในการแปลความหมายของวิดีโอที่ลำดับภาพ Sign I สูงขึ้น 0.57 ซึ่งแสดงให้เห็นว่าวิธีการที่นำเสนอช่วยปรับปรุงให้คุณภาพของวิดีโอที่มีความชัดเจนในการแปลความหมายเพิ่มขึ้น รูปที่ 4.15 และรูปที่ 4.16 กราฟค่าเฉลี่ย CIM ของลำดับภาพ Sign IV กรณีเฟตติงช้าและของลำดับภาพ Sign I กรณีเฟตติงเร็ว รูปที่ 4.17 ถึง รูปที่ 4.20 แสดงคุณภาพของวิดีโอที่ลำดับภาพ Sign IV และ Sign I ที่ถอดรหัส



รูปที่ 4.15 กราฟค่าเฉลี่ย CIM เปรียบเทียบระหว่างวิธี FMO_Imp+Intra_Scheme, FMO_Dis+Intra_Scheme และ Intra_Scheme ของลำดับภาพ Sign IV กรณีเฟดดิ้งช้า



รูปที่ 4.16 กราฟค่าเฉลี่ย CIM เปรียบเทียบระหว่างวิธี FMO_Imp+Intra_Scheme, FMO_Dis+Intra_Scheme และ Intra_Scheme ของลำดับภาพ Sign I กรณีเฟดดิ้งเร็ว



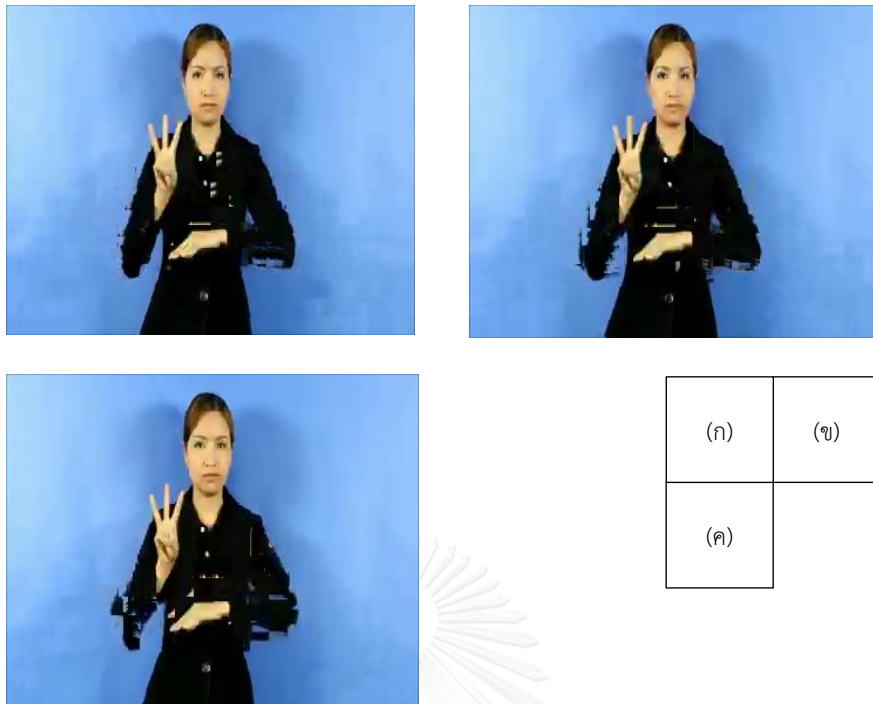
(ก)	(ข)
(ค)	

รูปที่ 4.17 เฟรมที่ 41 ของลำดับภาพ Sign IV ที่อัตราบิต 128 kbps กรณีเฟดดิ้งช้าโดย
(ก) FMO_Imp+Intra_Scheme (ข) FMO_Dis+Intra_Scheme (ค) Intra_Scheme

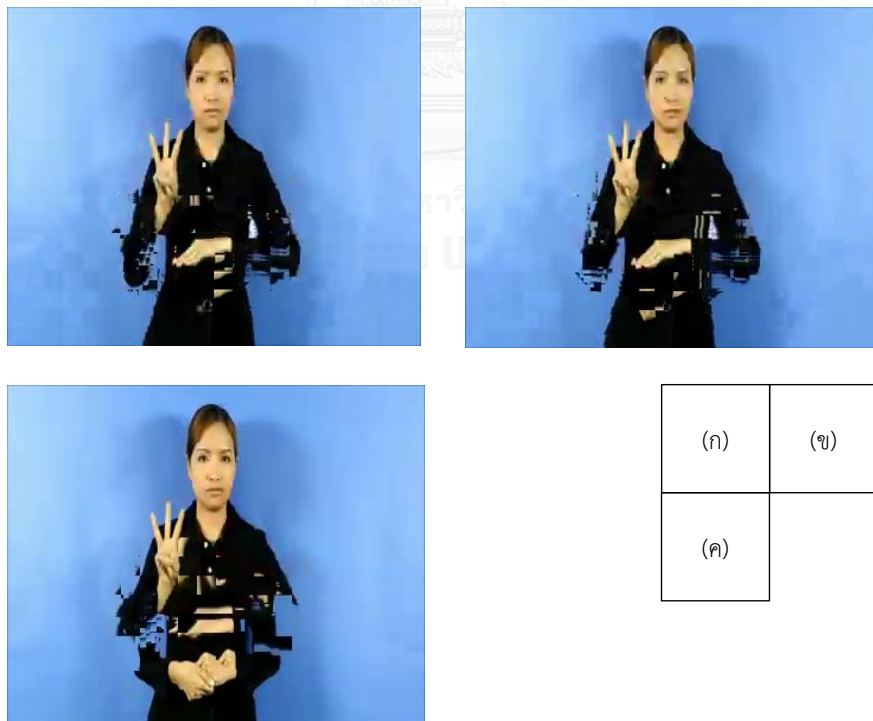


(ก)	(ข)
(ค)	

รูปที่ 4.18 เฟรมที่ 74 ของลำดับภาพ Sign IV ที่อัตราบิต 128 kbps กรณีเฟดดิ้งเร็วโดย
(ก) FMO_Imp+Intra_Scheme (ข) FMO_Dis+Intra_Scheme (ค) Intra_Scheme



รูปที่ 4.19 เฟรมที่ 54 ของลำดับภาพ Sign I ที่อัตราบิต 96 kbps กรณีเฟดดิ้งช้าโดย
(ก) FMO_Imp+Intra_Scheme (ข) FMO_Dis+Intra_Scheme (ค) Intra_Scheme



รูปที่ 4.20 เฟรมที่ 57 ของลำดับภาพ Sign I ที่อัตราบิต 96 kbps กรณีเฟดดิ้งเร็วโดย
(ก) FMO_Imp+Intra_Scheme (ข) FMO_Dis+Intra_Scheme (ค) Intra_Scheme

4.3.5 การปรับปรุงการควบคุมอัตราระดับเฟรมสำหรับมาตรฐานการเข้ารหัส H.264/AVC

4.3.5.1 แบบจำลองบิตส่วนหัว

ในหัวข้อที่ 3.3.1 ได้นำเสนอแบบจำลองบิตส่วนหัวเมื่อมีการใช้เครื่องมือต้านทานความผิดพลาดตามที่ได้นำเสนอ เพื่อเป็นการเปรียบเทียบความถูกต้องระหว่างค่าของแบบจำลองบิตส่วนหัวที่ได้จากการประมาณตามสมการที่ (3.6) และค่าจริงที่ได้จากการทดลอง สามารถคำนวณได้จากค่าสัมประสิทธิ์การตัดสินใจ (Coefficient of Determination : R^2) เมื่อค่าสัมประสิทธิ์การตัดสินใจมีค่าใกล้เคียง 1 แสดงว่าจำนวนบิตส่วนหัวที่ได้จากแบบจำลองมีค่าใกล้กับจำนวนบิตจริงที่ได้จากการทดลอง ตารางที่ 4.11 ถึงตารางที่ 4.14 แสดงค่า R^2 ของลำดับภาพ Sign I, Sign II, Sign III, Sign IV เปรียบเทียบระหว่างแบบจำลองบิตส่วนหัวที่ได้นำเสนอตามสมการที่ (3.6) (Header_FMOIntra) แบบจำลองบิตส่วนหัวใน [32] ตามสมการที่ (3.4) (Header_FMO) และแบบจำลองบิตส่วนหัวใน [35] ตามสมการที่ (3.5) (Header_NoFMO) เมื่อมีการใช้การจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้จำนวน 8 กลุ่มสไลซ์ต่อเฟรม ใช้การควบคุมอัตราที่แตกต่างกันและภายใต้การจำลองความผิดพลาดจากช่องสัญญาณที่กรณีเฟตติงช้าและกรณีเฟตติงเร็ว

ตารางที่ 4.11 ค่า R^2 ที่อัตราบิต 64 kbps

ลำดับภาพ	R^2					
	Header_FMOIntra		Header_FMO		Header_NoFMO	
	เฟตติงช้า	เฟตติงเร็ว	เฟตติงช้า	เฟตติงเร็ว	เฟตติงช้า	เฟตติงเร็ว
Sign I	0.955	0.934	-2.570	-3.624	-0.654	-0.607
Sign II	0.896	0.916	-2.389	-2.913	-0.457	-0.794
Sign III	0.877	0.887	-0.547	-1.564	0.575	0.053
Sign IV	0.913	0.894	-1.687	-3.354	-0.113	-0.643

ตารางที่ 4.12 ค่า R^2 ที่อัตราบิต 96 kbps

ลำดับภาพ	R^2					
	Header_FMOIntra		Header_FMO		Header_NoFMO	
	เฟตติงช้า	เฟตติงเร็ว	เฟตติงช้า	เฟตติงเร็ว	เฟตติงช้า	เฟตติงเร็ว
Sign I	0.949	0.926	-1.900	-3.515	-0.103	-0.405
Sign II	0.912	0.923	-2.338	-2.108	0.052	-0.210
Sign III	0.921	0.863	0.063	-1.172	0.791	0.208
Sign IV	0.939	0.915	-1.514	-1.884	0.334	-0.073

ตารางที่ 4.13 ค่า R^2 ที่อัตราบิต 128 kbps

ลำดับภาพ	R^2					
	Header_FMOIntra		Header_FMO		Header_NoFMO	
	เฟตติงช้า	เฟตติงเร็ว	เฟตติงช้า	เฟตติงเร็ว	เฟตติงช้า	เฟตติงเร็ว
Sign I	0.969	0.958	-1.840	-1.884	0.173	0.197
Sign II	0.934	0.942	-1.602	-1.679	0.365	0.083
Sign III	0.927	0.918	0.141	-1.335	0.693	0.448
Sign IV	0.959	0.939	-0.916	-0.717	0.434	0.311

ตารางที่ 4.14 ค่า R^2 ที่อัตราบิต 160 kbps

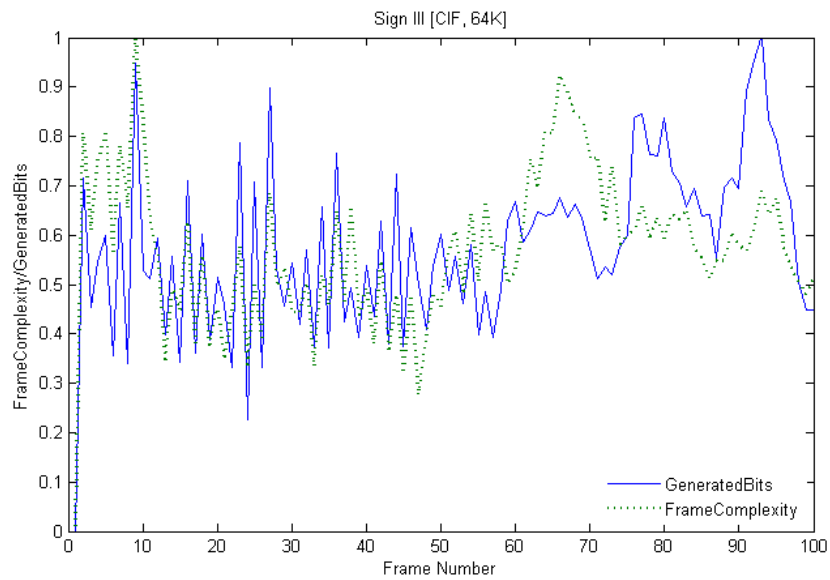
ลำดับภาพ	R^2					
	Header_FMOIntra		Header_FMO		Header_NoFMO	
	เฟตติงช้า	เฟตติงเร็ว	เฟตติงช้า	เฟตติงเร็ว	เฟตติงช้า	เฟตติงเร็ว
Sign I	0.980	0.978	-0.390	-0.715	0.582	0.518
Sign II	0.950	0.957	-0.703	-1.168	0.570	0.273
Sign III	0.956	0.961	0.173	-0.206	0.735	0.652
Sign IV	0.966	0.970	-0.432	-0.496	0.546	0.441

จากผลการทดลองพบว่าแบบจำลองบิตส่วนหัวที่ได้นำเสนอในวิทยานิพนธ์นี้ให้ความแม่นยำในการทำนายบิตส่วนหัวสูง คือมีค่า R^2 ใกล้เคียง 1 ในทุกอัตราบิตและการจำลองความผิดพลาด

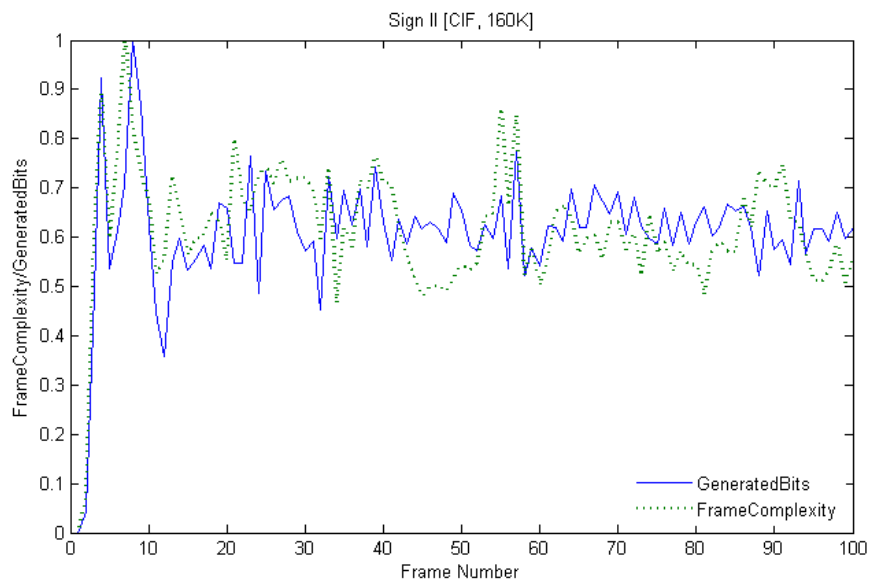
4.3.5.2 การวัดความซับซ้อนภายในเฟรม

ในหัวข้อที่ 3.3.1 ได้ปรับปรุงการควบคุมอัตราระดับเฟรมให้ดีขึ้นและเหมาะสมสำหรับวีดิทัศน์ภาษามือ โดยปรับปรุงการประมาณบิตเป้าหมายให้มีความสอดคล้องกับความซับซ้อนภายในเฟรม โดยใช้แบบจำลองความซับซ้อนภายในเฟรมตามสมการที่ (3.10) การเปรียบเทียบความสัมพันธ์ระหว่างบิตที่สร้างขึ้นจริงกับค่าจากการประมาณความซับซ้อนภายในเฟรมที่ได้จากแบบจำลองถูกนำมาใช้เพื่อเป็นการเปรียบเทียบความถูกต้องของแบบจำลอง เมื่อมีการใช้การจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้จำนวน 8 กลุ่มสไลซ์ต่อเฟรม รูปที่ 4.21 และรูปที่ 4.22 เป็นการเปรียบเทียบระหว่างบิตที่สร้างขึ้นจริงเมื่อทำการเข้ารหัสเทียบกับค่าจากการประมาณความซับซ้อนภายในเฟรมที่ถูกระบุให้ (Normalize) ของลำดับภาพ Sign III ที่อัตราบิต 64 kbps และ Sign II ที่อัตราบิต 160 kbps

ตามลำดับจากรูปแสดงให้เห็นว่าค่าความซับซ้อนภายในเฟรมมีความสัมพันธ์กันสูงกับจำนวนของบิตที่สร้างขึ้นจากการเข้ารหัส



รูปที่ 4.21 เปรียบเทียบระหว่างบิตที่สร้างขึ้นจริงกับค่าจากการประมาณความซับซ้อนภายในเฟรมของลำดับภาพ Sign III ที่อัตราบิต 64 kbps



รูปที่ 4.22 เปรียบเทียบระหว่างบิตที่สร้างขึ้นจริงกับค่าจากการประมาณความซับซ้อนภายในเฟรมของลำดับภาพ Sign II ที่อัตราบิต 160 kbps

ผลการทดลองของการปรับปรุงการควบคุมอัตราการระดับเฟรมแสดงในตารางที่ 4.15 และ ตารางที่ 4.16 เปรียบเทียบค่าเฉลี่ย CIM ของวิดิทัศน์เมื่อมีการใช้เครื่องมือด้านทานความผิดพลาดที่ได้นำเสนอและการควบคุมอัตราบิตระดับเฟรม ($Er_{tool}+Fr_{RC}$) กับเครื่องมือด้านทานความผิดพลาดที่ได้นำเสนอโดยไม่มีการใช้การควบคุมอัตราบิตระดับเฟรม (Er_{tool}) ภายใต้การจำลองความผิดพลาดจากช่องสัญญาณกรณีเฟตติงช้าและกรณีเฟตติงเร็ว จากผลการทดลองพบว่าความชัดเจนในการแปลความหมายของวิดิทัศน์ลดลงเกือบทุกกรณีเนื่องจากบิตเป้าหมายสำหรับการเข้ารหัสในแต่ละเฟรมถูกหักลบด้วยค่าบิตส่วนหัวที่คำนวณได้

ตารางที่ 4.15 เปรียบเทียบค่าเฉลี่ย CIM ในบริเวณของผู้ทำท่าทางามือกรณีเฟตติงช้า

เฟตติงช้า								
ลำดับภาพ วิธี	64 kbps				96 kbps			
	Sign I	Sign II	Sign III	Sign IV	Sign I	Sign II	Sign III	Sign IV
$Er_{tool}+Fr_{RC}$	2.29	2.29	2.41	2.28	2.36	2.36	2.41	2.39
Er_{tool}	2.33	2.33	2.42	2.35	2.33	2.41	2.41	2.40
ลำดับภาพ วิธี	128 kbps				160 kbps			
	Sign I	Sign II	Sign III	Sign IV	Sign I	Sign II	Sign III	Sign IV
$Er_{tool}+Fr_{RC}$	2.07	2.07	2.06	2.22	2.10	2.10	1.95	2.27
Er_{tool}	2.11	2.29	2.07	2.23	2.13	2.34	1.90	2.26

ตารางที่ 4.16 เปรียบเทียบค่าเฉลี่ย CIM ในบริเวณของผู้ทำท่าทางามือกรณีเฟตติงเร็ว

เฟตติงเร็ว								
ลำดับภาพ วิธี	64 kbps				96 kbps			
	Sign I	Sign II	Sign III	Sign IV	Sign I	Sign II	Sign III	Sign IV
$Er_{tool}+Fr_{RC}$	2.04	2.04	1.85	2.05	2.05	2.05	1.74	2.11
Er_{tool}	2.03	1.96	1.86	2.11	2.00	1.78	1.77	2.13
ลำดับภาพ วิธี	128 kbps				160 kbps			
	Sign I	Sign II	Sign III	Sign IV	Sign I	Sign II	Sign III	Sign IV
$Er_{tool}+Fr_{RC}$	1.96	1.96	1.78	2.21	1.91	1.91	1.75	2.16
Er_{tool}	1.96	1.76	1.77	2.14	1.91	1.68	1.77	2.22

4.3.6 การปรับปรุงการควบคุมอัตราการระดับหน่วยพื้นฐานย่อยสำหรับมาตรฐานการเข้ารหัส H.264/AVC

การควบคุมอัตราการระดับหน่วยพื้นฐานย่อยถูกนำมาใช้ในการปรับปรุงคุณภาพของมาโครบล็อก บริเวณใบหน้า มือและลำตัวของผู้ทำท่าภาษามือตามแบบจำลองทางภาษาที่คุณภาพของวีดิทัศน์ ลดลงจากการพิจารณาผลกระทบของบิตส่วนหัว ตารางที่ 4.17 และ ตารางที่ 4.18 เปรียบเทียบ ค่าเฉลี่ย CIM ของวีดิทัศน์ระหว่าง 3 วิธีคือ การใช้เครื่องมือด้านทานความผิดพลาดที่ได้นำเสนอและ การควบคุมอัตราบิตระดับเฟรมและหน่วยพื้นฐานย่อย (Er_tool+Fr_MB_RC), การใช้เครื่องมือ ด้านทานความผิดพลาดที่ได้นำเสนอและการควบคุมอัตราการระดับเฟรม (Er_tool+Fr_RC) และวิธี สุดท้ายคือวิธี AAIR ที่ใช้มาโครบล็อกภายในในการเข้ารหัสทุกมาโครบล็อกของใบหน้า มือและลำตัว ทุกๆ 3 เฟรมภายใต้การจำลองความผิดพลาดจากช่องสัญญาณกรณีเฟตดิงช้าและกรณีเฟตดิงเร็ว

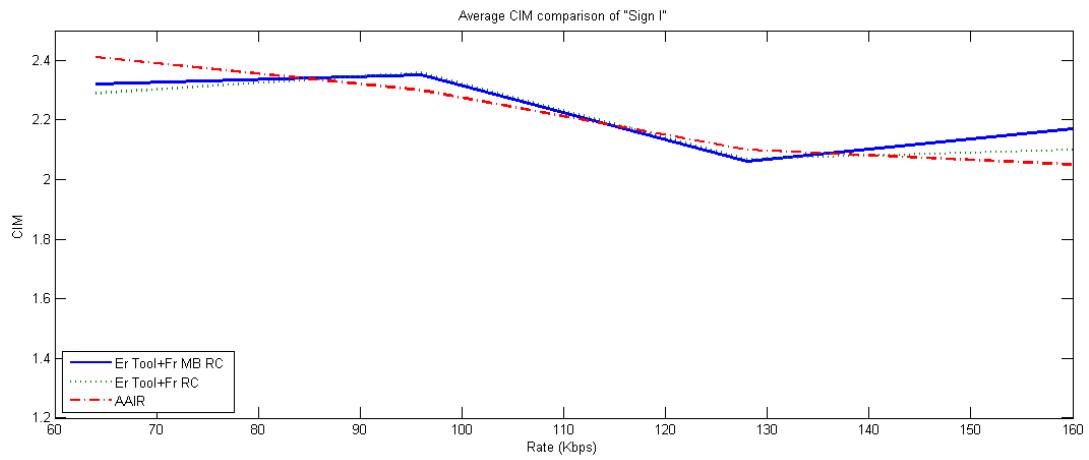
ตารางที่ 4.17 เปรียบเทียบค่าเฉลี่ย CIM ในบริเวณของผู้ทำท่าภาษามือกรณีเฟตดิงช้า

เฟตดิงช้า								
ลำดับภาพ วิธี	64 kbps				96 kbps			
	Sign I	Sign II	Sign III	Sign IV	Sign I	Sign II	Sign III	Sign IV
Er_tool+ Fr_MB_RC	2.32	2.42	2.45	2.32	2.35	2.45	2.45	2.44
Er_tool+Fr_RC	2.29	2.36	2.41	2.28	2.36	2.43	2.41	2.39
AAIR	2.41	2.41	2.03	2.29	2.30	2.69	2.30	2.25
ลำดับภาพ วิธี	128 kbps				160 kbps			
	Sign I	Sign II	Sign III	Sign IV	Sign I	Sign II	Sign III	Sign IV
Er_tool+ Fr_MB_RC	2.06	2.26	2.06	2.24	2.17	2.36	1.93	2.25
Er_tool+Fr_RC	2.07	2.30	2.06	2.22	2.10	2.34	1.95	2.27
AAIR	2.10	1.89	1.82	2.22	2.05	1.93	1.73	2.00

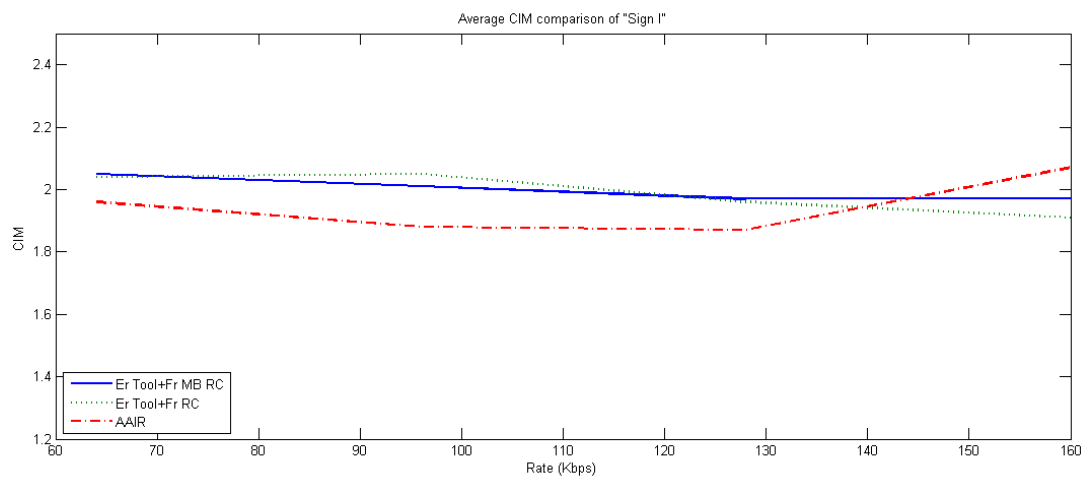
ตารางที่ 4.18 เปรียบเทียบค่าเฉลี่ย CIM ในบริเวณของผู้ทำท่าภาษามือกรณีเฟตติงเร็ว

เฟตติงเร็ว									
วิธี	ลำดับภาพ	64 kbps				96 kbps			
		Sign I	Sign II	Sign III	Sign IV	Sign I	Sign II	Sign III	Sign IV
Er_tool+ Fr_MB_RC		2.05	2.00	1.81	2.08	2.01	1.86	1.73	2.14
Er_tool+Fr_RC		2.04	1.91	1.85	2.05	2.05	1.83	1.74	2.11
AAIR		1.96	1.74	1.65	1.98	1.88	1.60	1.61	2.04
วิธี	ลำดับภาพ	128 kbps				160 kbps			
		Sign I	Sign II	Sign III	Sign IV	Sign I	Sign II	Sign III	Sign IV
Er_tool+ Fr_MB_RC		1.97	1.72	1.79	2.24	1.97	1.67	1.80	2.16
Er_tool+Fr_RC		1.96	1.76	1.78	2.21	1.91	1.68	1.75	2.16
AAIR		1.87	1.51	1.66	2.02	2.07	1.70	1.72	2.11

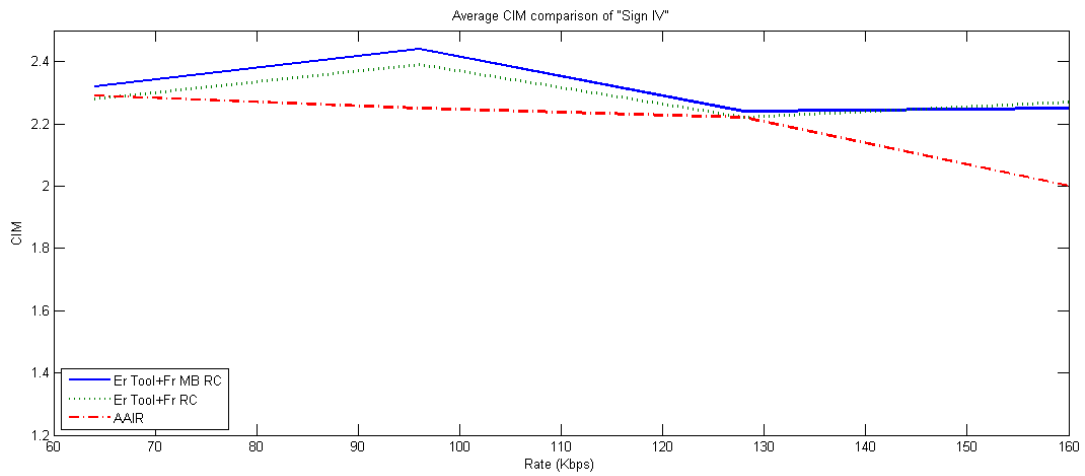
จากผลการทดลองเมื่อทำการเปรียบเทียบการใช้เครื่องมือด้านทานความผิดพลาดที่ได้ นำเสนอและการควบคุมอัตราบิดระดับเฟรมและหน่วยพื้นฐานย่อย (Er_tool+Fr_MB_RC) กับการใช้ เครื่องมือด้านทานความผิดพลาดที่ได้ นำเสนอและการควบคุมอัตราบิดระดับเฟรม (Er_tool+Fr_RC) พบว่าการเข้ารหัสที่อัตราบิด 160 kbps ความชัดเจนในการแปลความหมายของวิดีโอที่ลำดับภาพ Sign I สูงขึ้น 0.07 และที่อัตราบิด 160 kbps ความชัดเจนในการแปลความหมายของลำดับภาพ วิดีโอที่ Sign I สูงขึ้น 0.06 ทั้งในกรณีเฟตติงช้าและเฟตติงเร็วตามลำดับ และเมื่อเทียบกับวิธี AAIR พบว่าการเข้ารหัสที่อัตราบิด 160 kbps ความชัดเจนในการแปลความหมายของวิดีโอที่ลำดับภาพ Sign II สูงขึ้น 0.43 และที่อัตราบิด 96 kbps ความชัดเจนในการแปลความหมายของลำดับภาพ วิดีโอที่ Sign II สูงขึ้น 0.26 ทั้งในกรณีเฟตติงช้าและเฟตติงเร็วตามลำดับ ผลการทดลองแสดงให้เห็นว่า การนำการควบคุมอัตราบิดระดับหน่วยพื้นฐานย่อยมาใช้ช่วยให้ความชัดเจนในการแปลความหมายของ วิดีโอที่ลำดับภาพมือทั้งกรณีเฟตติงช้าและเฟตติงเร็วเพิ่มขึ้น กราฟค่าเฉลี่ย CIM ของลำดับภาพ Sign I กรณีเฟตติงช้าและเร็วแสดงตามรูปที่ 4.23 และรูปที่ 4.24 ตามลำดับ ส่วนรูปที่ 4.25 และรูปที่ 4.26 กราฟค่าเฉลี่ย CIM ของลำดับภาพ Sign IV กรณีเฟตติงช้าและเร็วตามลำดับ แสดงคุณภาพของวิดีโอที่ลำดับภาพ Sign II และ Sign III ที่ถอดรหัส



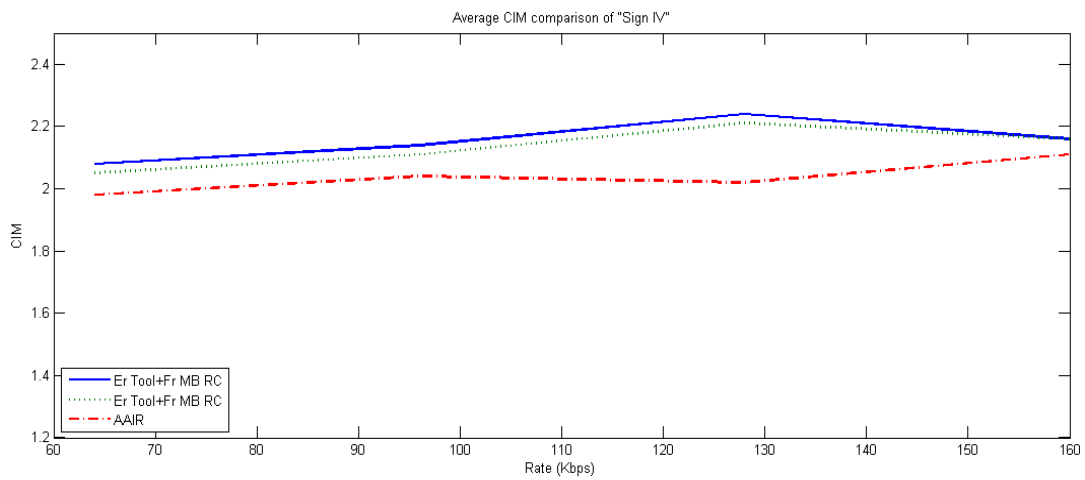
รูปที่ 4.23 กราฟค่าเฉลี่ย CIM เปรียบเทียบระหว่างวิธี Er_Tool+Fr_MB_RC, Er_Tool+Fr_RC และ AAIR ของลำดับภาพ Sign I กรณีเฟตดิ่งช้า



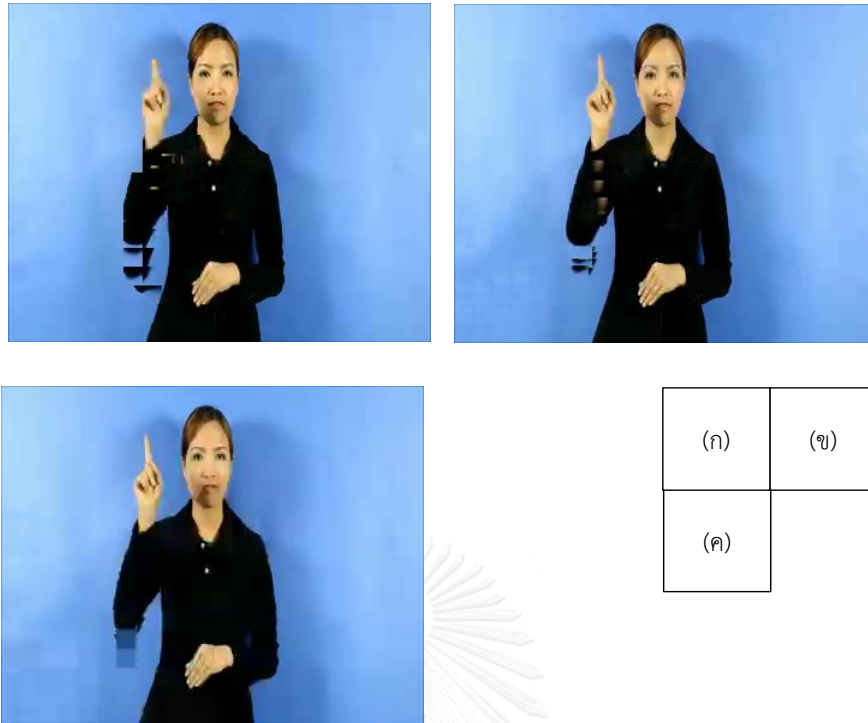
รูปที่ 4.24 กราฟค่าเฉลี่ย CIM เปรียบเทียบระหว่างวิธี Er_Tool+Fr_MB_RC, Er_Tool+Fr_RC และ AAIR ของลำดับภาพ Sign I กรณีเฟตดิ่งเร็ว



รูปที่ 4.25 กราฟค่าเฉลี่ย CIM เปรียบเทียบระหว่างวิธี Er_Tool+Fr_MB_RC, Er_Tool+Fr_RC และ AAIR ของลำดับภาพ Sign IV กรณีเฟดดิ้งช้า

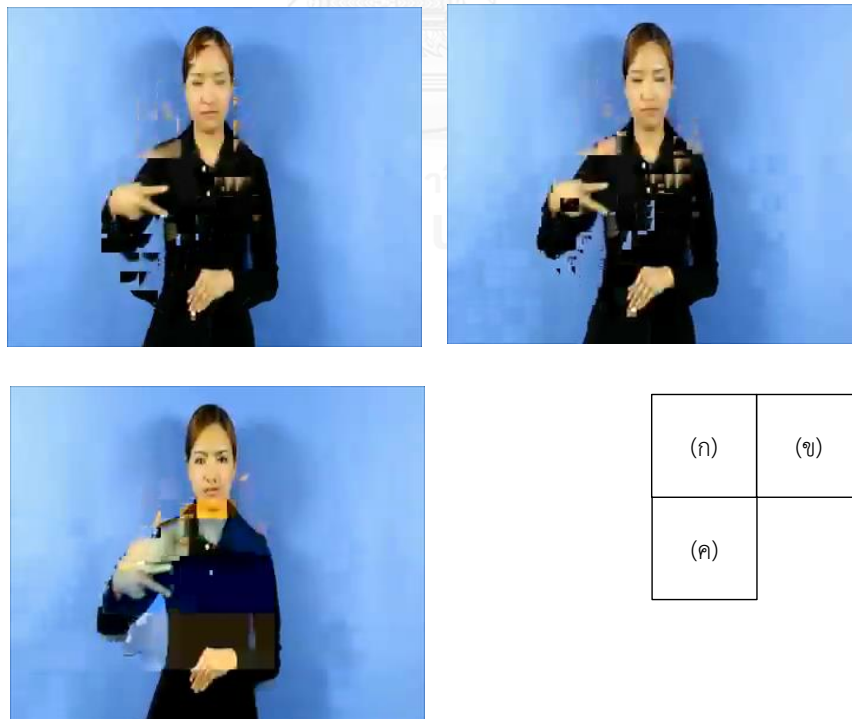


รูปที่ 4.26 กราฟค่าเฉลี่ย CIM เปรียบเทียบระหว่างวิธี Er_Tool+Fr_MB_RC, Er_Tool+Fr_RC และ AAIR ของลำดับภาพ Sign IV กรณีเฟดดิ้งเร็ว



รูปที่ 4.27 เฟรมที่ 16 ของลำดับภาพ Sign II ที่อัตราบิต 64 kbps กรณีเฟตติงช้าโดย

(ก) Er_tool+Fr_MB_RC (ข) Er_tool+Fr_RC (ค) AAIR



รูปที่ 4.28 เฟรมที่ 71 ของลำดับภาพ Sign II ที่อัตราบิต 64 kbps กรณีเฟตติงเร็วโดย

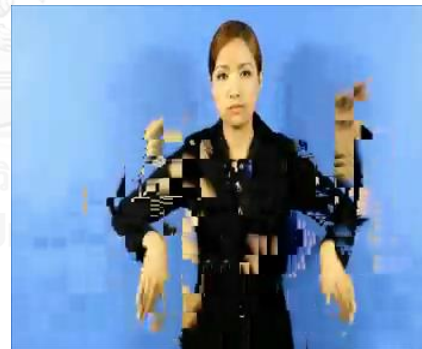
(ก) Er_tool+Fr_MB_RC (ข) Er_tool+Fr_RC (ค) AAIR



(ก)	(ข)
(ค)	

รูปที่ 4.29 เฟรมที่ 87 ของลำดับภาพ Sign III ที่อัตราบิต 128 kbps กรณีเฟดดิ้งช้าโดย

(ก) Er_tool+Fr_MB_RC (ข) Er_tool+Fr_RC (ค) AAIR



(ก)	(ข)
(ค)	

รูปที่ 4.30 เฟรมที่ 85 ของลำดับภาพ Sign III ที่อัตราบิต 128 kbps กรณีเฟดดิ้งเร็วโดย

(ก) Er_tool+Fr_MB_RC (ข) Er_tool+Fr_RC (ค) AAIR

บทที่ 5

สรุปผลการจำลองและข้อเสนอแนะ

เนื้อหาในบทนี้เป็นสรุปผลการจำลอง และข้อเสนอแนะสำหรับงานวิจัยในอนาคตที่ต้องการพัฒนาประสิทธิภาพเทคนิคการต้านทานความผิดพลาดสำหรับการส่งวีดิทัศน์ภาษามือ

5.1 สรุปผลการจำลอง

วิทยานิพนธ์นี้นำเสนอการพัฒนาเทคนิคโดยใช้เครื่องมือต้านทานความผิดพลาดในมาตรฐาน H.264 และการควบคุมอัตราบิตโดยพิจารณาร่วมกับแบบจำลองทางภาษาเพื่อลดผลกระทบของวีดิทัศน์จากการสูญหายของข้อมูล เครื่องมือต้านทานความผิดพลาดที่ใช้ได้แก่การคืนสภาพมาโครบล็อกภายในและการจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้ แผนการคืนสภาพมาโครบล็อกภายในถูกนำเสนอเพื่อลดผลของการแพร่กระจายความผิดพลาดเชิงเวลาที่เกิดขึ้นในเฟรมก่อนหน้าจากช่องสัญญาณซึ่งเป็นปัจจัยหลักที่ทำให้คุณภาพของวีดิทัศน์ลดลง โดยการจัดสรรมาโครบล็อกภายในบริเวณใบหน้า มือ และลำตัวของผู้ทำภาษามือซึ่งมีความสำคัญต่อการแปลความหมายได้ถูกต้อง แบบจำลองอัตราคืนสภาพมาโครบล็อกภายในถูกนำมาใช้ในการหาจำนวนของมาโครบล็อกภายในที่เหมาะสมโดยพิจารณาจากข้อมูลป้อนกลับและบิตเป้าหมายที่ใช้ในการเข้ารหัสในแต่ละเฟรมเพื่อไม่ให้กระทบกับจำนวนบิตที่ใช้ในการเข้ารหัสเฟรมถัดไป ส่วนในบริเวณลำตัวจะจัดสรรมาโครบล็อกภายในเพื่อคืนสภาพวีดิทัศน์ในบริเวณที่มือพาดผ่านเพื่อลดผลกระทบของการปกปิดความผิดพลาดแบบไม่มีการชดเชยการเคลื่อนที่ ผลของการจำลองแสดงให้เห็นว่าวิธีที่นำเสนอช่วยเพิ่มค่า PSNR ให้สูงขึ้นทั้งกรณีเฟตติงช้าและเฟตติงเร็ว

การจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้เป็นวิธีที่ช่วยยับยั้งการแพร่กระจายความผิดพลาดเชิงพื้นที่ภายในเฟรมนอกเหนือจากวิธีคืนสภาพมาโครบล็อกภายใน ถึงแม้ว่าจะมีหลายงานวิจัยที่นำเสนอเกี่ยวกับการจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้ แต่การจัดสรรมาโครบล็อกเข้าสู่กลุ่มสไลซ์เพื่อให้มาโครบล็อกที่สำคัญได้รับผลกระทบจากความผิดพลาดน้อยลงยังคงเป็นประเด็นที่ท้าทาย ในงานนี้ได้ปรับวิธีการจากการจัดสรรมาโครบล็อกที่พิจารณาเพียงแค่ว่าได้ความเพี้ยนในแต่ละกลุ่มสไลซ์ใกล้เคียงกันมาเป็นการพิจารณาความเพี้ยนในบริเวณใบหน้า มือและลำตัวให้ใกล้เคียงกันในแต่ละกลุ่มสไลซ์เพื่อรักษาคุณภาพของวีดิทัศน์ให้ง่ายต่อการแปลความหมายได้ถูกต้อง จากผลการจำลองแสดงให้เห็นว่าวิธีที่นำเสนอสามารถจัดให้ค่าความเพี้ยนของหน้า มือและลำตัวในแต่ละกลุ่มสไลซ์มีค่าใกล้เคียงกันมากกว่าการจัดแบบเดิมทั้งกรณีเฟตติงช้าและเฟตติงเร็วโดยพิจารณาจากค่าเบี่ยงเบน

มาตรฐานและยังช่วยให้ค่า CIM สูงขึ้นในหลายกรณีซึ่งทำให้วิดิทัศน์ภาษามือมีความชัดเจนในการแปลความหมายเพิ่มขึ้น

การควบคุมอัตราที่นำมาใช้งานมีทั้งระดับเฟรมและหน่วยพื้นฐานย่อย ในการควบคุมอัตราระดับเฟรมขั้นต้นแรกมีการปรับปรุงแบบจำลองบิตส่วนหัวให้มีความแม่นยำมากขึ้นเมื่อมีการใช้การจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้และการคืนสภาพมาโครบล็อกภายใน จากผลการเปรียบเทียบค่าบิตส่วนหัวที่ได้จากการคำนวณโดยใช้แบบจำลองกับค่าบิตส่วนหัวจริงที่ได้จากการทดลองพบว่าค่าบิตส่วนหัวที่ได้จากแบบจำลองให้ค่าที่ใกล้เคียงกับค่าจริงมากทั้งกรณีเฟตติงช้าและเฟตติงเร็ว เนื่องจากแบบจำลองถูกปรับปรุงให้มีการพิจารณาค่าบิตส่วนหัวที่เกิดขึ้นจากมาโครบล็อกภายใน ขั้นตอนที่สองเพิ่มประสิทธิภาพของการเข้ารหัสโดยปรับปรุงการประมาณบิตเป้าหมายให้มีความสอดคล้องกับแบบจำลองการวัดความซับซ้อนภายในเฟรม การเปรียบเทียบความสัมพันธ์ระหว่างบิตที่สร้างขึ้นจริงกับค่าจากการประมาณความซับซ้อนภายในเฟรมที่ได้จากแบบจำลองแสดงให้เห็นว่าค่าที่ได้จากแบบจำลองมีความถูกต้องในการประมาณความซับซ้อนภายในเฟรม บิตเป้าหมายที่ถูกปรับปรุงด้วยแบบจำลองการวัดความซับซ้อนภายในเฟรมจะถูกหักลบด้วยค่าบิตส่วนหัวที่คำนวณได้ในตอนแรก เพื่อให้ได้บิตเป้าหมายตรงตามที่ตัวเข้ารหัสกำหนดแต่ในขณะเดียวกันคุณภาพของวิดิทัศน์หลังจากใช้การควบคุมอัตราที่ระดับเฟรมค่า CIM จะลดต่ำลงในเกือบทุกกรณีเนื่องจากค่าบิตเป้าหมายที่ใช้ในการเข้ารหัสแต่ละเฟรมลดลงจากการพิจารณาผลกระทบของบิตส่วนหัว คุณภาพวิดิทัศน์ที่ลดลงโดยเฉพาะในบริเวณใบหน้าและมือของผู้ทำท่าภาษามือส่งผลต่อการแปลความหมายที่ต้องภายในวิดิทัศน์ทำได้ยากขึ้น การควบคุมอัตราระดับหน่วยพื้นฐานย่อยจึงถูกนำมาใช้ในการปรับปรุงคุณภาพของมาโครบล็อกบริเวณใบหน้า มือและลำตัวของผู้ทำท่าภาษามือตามแบบจำลองทางภาษา จากผลการจำลองแสดงให้เห็นว่าวิธีการควบคุมอัตราระดับหน่วยพื้นฐานย่อยช่วยให้คุณภาพของวิดิทัศน์ในด้านความชัดเจนในการแปลความหมายสูงขึ้นทั้งกรณีเฟตติงช้าและเฟตติงเร็ว เนื่องจากมีการลดบิตที่ใช้ในการเข้ารหัสบริเวณพื้นหลังและจัดสรรบิตในส่วนที่ลดได้นี้ให้กับบริเวณใบหน้า มือและลำตัวตามลำดับความสำคัญ

5.2 ข้อเสนอแนะสำหรับงานวิจัยในอนาคต

ในวิทยานิพนธ์นี้พัฒนาเครื่องมือด้านทานความผิดพลาดอันได้แก่การคืนสภาพมาโครบล็อกภายในและการจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้ร่วมกับการควบคุมอัตราสำหรับวิดิทัศน์ภาษามือสามารถปรับปรุงให้มีประสิทธิภาพมากขึ้นได้ดังข้อต่อไปนี้

1. อัตราคืนสภาพมาโครบล็อกภายในที่เหมาะสมถูกคำนวณโดยใช้แบบจำลองซึ่งแบบจำลองนั้นถูกสร้างขึ้นโดยใช้ค่ากลุ่มตัวอย่างที่ได้จากการทดลองอย่างละเอียดที่ความผิดพลาดของช่องสัญญาณในแต่ละความถี่เพียง 1 ครั้ง เพื่อให้ได้แบบจำลองที่มีประสิทธิภาพสูงสุดจึงควรทำการทดสอบมากกว่า 1 ครั้งเพื่อให้ได้ค่าที่เหมาะสม
2. แบบจำลองอัตราคืนสภาพมาโครบล็อกภายในใช้ในการหาจำนวนมาโครบล็อกภายในที่เหมาะสมต่อเฟรม เมื่อมีการใช้คู่กับการจัดเรียงมาโครบล็อกแบบยืดหยุ่นได้สามารถปรับเป็นการหาจำนวนมาโครบล็อกภายในต่อเฟรมที่เหมาะสมซึ่งอาจจะช่วยให้คุณภาพของวิดิทัศน์เพิ่มขึ้นได้
3. สถานะของช่องสัญญาณและเทคนิคการควบคุมความผิดพลาดควรมีการพิจารณาร่วมด้วยเพื่อช่วยให้ด้านทานต่อความผิดพลาดที่จะเกิดขึ้นในวิดิทัศน์ได้มากขึ้น
4. เพื่อปรับปรุงคุณภาพของวิดิทัศน์ให้ดีขึ้นการปกปิดความผิดพลาดแบบขดเชยการเคลื่อนที่จึงควรนำมาใช้งาน

รายการอ้างอิง

- [1] N. Cherniavsky, A. C. Cavender, R. E. Ladner *et al.*, "Variable frame rate for low power mobile sign language communication." pp. 163-170.
- [2] S. Hooper, C. Miller, S. Rose *et al.*, "The Effects of Digital Video Quality on Learner Comprehension in an American Sign Language Assessment Environment," *Sign Language Studies* vol. 8, no. 1, 2007.
- [3] J. J. Tran, T. W. Johnson, J. Kim *et al.*, "A web-based user survey for evaluating power saving strategies for deaf users of mobileASL," in Proceedings of the 12th international ACM SIGACCESS conference on Computers and accessibility, Orlando, Florida, USA, 2010, pp. 115-122.
- [4] F. M. Ciaramello, and S. S. Hemami, "A Computational Intelligibility Model for Assessment and Compression of American Sign Language Video," *IEEE Trans. Image Process.*, vol. 20, no. 11, pp. 3014-3027, 2011.
- [5] F. M. Ciaramello, K. Jung, and S. S. Hemami, "Quality versus intelligibility: Studying human preferences for american sign language video." pp. 70-73.
- [6] C. Jaehong, S. Whittle, E. A. Riskin *et al.*, "Improving Compressed Video Sign Language Conversations in the Presence of Data Loss." pp. 383-392.
- [7] "Electronic Thai Sign Language Communication System Development Project," *Final Technical Report*: Chula Unisearch, Chulalongkorn University, December 2012.
- [8] T. Wiegand, G. J. Sullivan, G. Bjontegaard *et al.*, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits. Syst. Video Technol.*, vol. 13, no. 7, pp. 560-576, 2003.
- [9] T. Stockhammer, and M. M. Hannuksela, "H.264/AVC video for wireless transmission," *Wireless Communications, IEEE*, vol. 12, no. 4, pp. 6-13, 2005.
- [10] A. R. Noerpel, L. Yi-Bing, and H. Sherry, "PACS: Personal Access Communications System-a tutorial," *IEEE Personal Communications*, vol. 3, no. 3, pp. 32-43, 1996.

- [11] S. Aramvith, C.-W. Lin, S. Roy *et al.*, "Wireless video transport using conditional retransmission and low-delay interleaving," *IEEE Trans. Circuits Syst. Video Technology*, vol. 12, no. 6, pp. 558-565, 2002.
- [12] A. Dimitris, C. Nishan, R. B. David *et al.*, "A perceptually optimised video coding system for sign language communication at low bit rates," *Signal Processing-Image Communication*, vol. 21, pp. 531--549, 2006.
- [13] E. Karen, R. Thompson, and R. Colvin, "Eye Gaze During Comprehension of American Sign Language by Native and Beginning Signers," *Journal of Deaf Studies and Deaf Education*, vol. 14, pp. 237--243, 2008.
- [14] F. Ciaramello, K. Jung, and S. Hemami, "Quality versus intelligibility: Evaluating the coding trade-offs for American Sign Language video." pp. 1-5.
- [15] K. I. Nakazono, A., "Digital Encoding Applied to Sign Language Video," *IEICE Trans. Inf. Syst.*, vol. E89-D, no. 6, pp. 8, 2006.
- [16] M. Ries, "Video Quality Estimation for Mobile Video Streaming," Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology, 2008.
- [17] "Draft Application Profile Sign language and Lip-reading real time conversation application of low bitrate video communication," ITU - Telecommunications Standardization Sector, 1998.
- [18] A. Cavender, R. E. Ladner, and E. A. Riskin, "MobileASL: Intelligibility of Sign Language Video as Constrained by Mobile Phone Technology." pp. 71-78.
- [19] D. Agrafiotis, N. Canagarajah, D. R. Bull *et al.*, "A perceptually optimised video coding system for sign language communication at low bit rates," *Signal Process.: Image Commun.*, vol. 21, no. 7, pp. 531-549, 8//, 2006.
- [20] D. M. Saxe, and R. A. Foulds, "Robust region of interest coding for improved sign language telecommunication," *IEEE Trans. Inf. Technol. Biomed.*, vol. 6, no. 4, pp. 310-316, 2002.
- [21] R. P. Schumeyer, E. A. Heredia, and K. E. Barner, "Region of interest priority coding for sign language videoconferencing." pp. 531-536.
- [22] "MobileASL," <http://mobileasl.cs.washington.edu>.

- [23] J. J. Tran, J. Kim, J. Chon *et al.*, "Evaluating quality and comprehension of real-time sign language video on mobile phones," in Proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility, Dundee, Scotland, UK, 2011, pp. 115-122.
- [24] F. M. H. Ciaramello, S. S. , "“Can You See Me Now?’ An Objective Metric for Predicting Intelligibility of Compressed American Sign Language Video," in Proc. SPIE Human Vision and Electronic Imaging, San Jose, CA, 2007.
- [25] "H.264 JM Ref. Software," <http://iphome.hhi.de/suehring/tml/download/>.
- [26] C. Quqing, C. Zhibo, G. Xiaodong *et al.*, "Attention-based adaptive intra refresh for error-prone video transmission," *IEEE Commun. Mag.*, vol. 45, no. 1, pp. 52-60, 2007.
- [27] P. Nunes, L. D. Soares, and F. Pereira, "Error resilient macroblock rate control for H.264/AVC video coding." pp. 2132-2135.
- [28] P. Nunes, L. D. Soares, and F. Pereira, "Automatic and adaptive network-aware macroblock intra refresh for error-resilient H.264/AVC video coding." pp. 3073-3076.
- [29] L. D. Soares, P. Nunes, and F. Pereira, "Efficient network-aware macroblock mode decision for error resilient H.264/AVC video coding." pp. 70730U-70730U-12.
- [30] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the Society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 11, Jun., 1963.
- [31] R. D. Cajote, S. Aramvith, R. Guevara *et al.*, "Improved Sorting Algorithm for Explicit FMO Macroblock Classification in H.264." pp. 514-519.
- [32] R. Cajote, S. Aramvith, and Y. Miyanaga, "FMO-based H.264 frame layer rate control for low bit rate video transmission," *EURASIP J. on Advances in Signal Process.*, vol. 2011, no. 1, pp. 63, 2011.
- [33] K. Do-Kyoung, S. Mei-Yin, and C. C. J. Kuo, "Rate Control for H.264 Video With Enhanced Rate and Distortion Models," *IEEE Trans. Circuits. Syst. Video Technol.*, vol. 17, no. 5, pp. 517-529, 2007.

- [34] T. Suksil, and T. H. Chalidabhongse, “Hand detection and feature extraction for static Thai Sign Language recognition,” in Proceedings of the 7th International Conference on Ubiquitous Information Management and Communication, Kota Kinabalu, Malaysia, 2013, pp. 1-6.
- [35] D.-K. Kwon, M.-Y. Shen, and C.-C. J. Kuo, “Rate control for H.264 video with enhanced rate and distortion models,” *IEEE Trans. Circuits. Syst. Video Technol.*, vol. 17, no. 5, pp. 517 - 529, 2007.



ภาคผนวก



ประวัติผู้เขียนวิทยานิพนธ์

นายศิริวิชญ์ สุปงกช สำเร็จการศึกษาปริญญาวิศวกรรมศาสตรบัณฑิต สาขาวิชาวิศวกรรมไฟฟ้า จากมหาวิทยาลัยศรีนครินทรวิโรฒ ในปีการศึกษา 2552 และเข้าศึกษาต่อในหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต ในสังกัดห้องปฏิบัติการกรรมวิธีสัญญาณดิจิทัล ภาควิชาวิศวกรรมไฟฟ้า ที่จุฬาลงกรณ์มหาวิทยาลัยในภาคการศึกษาที่ 2 ปีการศึกษา 2553

