

บทที่ 1 บทนำ

1.1 ความสำคัญและความเป็นมาของปัญหา

ในปัจจุบันการส่งผ่านข้อมูลข่าวสารต่างๆ ในรูปแบบดิจิทัลได้รับความนิยมอย่างมาก โดยตัวข้อมูลข่าวสารจะถูกเปลี่ยนจากสัญญาณอนาล็อกเป็นดิจิทัล ก่อนจะถูกประมวลผลและส่งผ่านไปยังสถานที่ต่างๆ ที่ต้องการ ในการใช้งานข้อมูลข่าวสารเหล่านั้น ตัวข้อมูลที่มีอยู่จำเป็นต้องถูกบีบอัดให้มีขนาดเล็กลงเสียก่อน เพื่อเพิ่มประสิทธิภาพในการเก็บรักษา เช่น ใช้เนื้อที่ในการเก็บน้อยลง หรือเพิ่มประสิทธิภาพในการส่งผ่าน เช่น ใช้ความกว้าง (Bandwidth) ของการส่งสัญญาณน้อยลง นอกจากนี้การส่งผ่านข้อมูลข่าวสารผ่านทางเครือข่ายอินเทอร์เน็ตก็ยังเป็นที่นิยมอย่างมากในทุกๆ ระดับ ทำให้มีความต้องการในการพัฒนาเทคโนโลยีทางด้านที่เกี่ยวข้องเพิ่มขึ้นอย่างรวดเร็ว เทคโนโลยีที่มีการพัฒนาสามารถแบ่งออกเป็น 2 ด้านได้คือ

- เทคโนโลยีที่ใช้เพิ่มความกว้างของช่องทางการส่งสัญญาณให้มีขนาดสูงขึ้น เช่น การใช้สื่อสัญญาณในการส่งข้อมูลที่เหมาะสม ซึ่งรวมไปถึงการพัฒนาเทคโนโลยีต่าง ๆ ทางด้านระบบการสื่อสารข้อมูล
- เทคโนโลยีทางด้านการประมวลผลข้อมูลให้มีความเร็ว และประสิทธิภาพสูงขึ้น เช่น การคิดค้นวิธีการบีบอัดข้อมูลให้มีขนาดเล็กกว่าวิธีการที่มีอยู่ในปัจจุบัน ทำให้การเก็บรักษา และส่งผ่านข้อมูลมีประสิทธิภาพสูงขึ้น

ประเทศต่าง ๆ ทั่วโลกได้ให้ความสนใจกับการพัฒนาเทคโนโลยีดังกล่าวข้างต้น เช่น การวางระบบเครือข่ายใยแก้วนำแสง และการพัฒนาวิธีการที่ใช้ในการประมวลผลข้อมูลอยู่ตลอดเวลา ภายในเวลาไม่กี่ปีที่ผ่านมา เราจะสังเกตเห็นว่ามีซอฟต์แวร์ ฮาร์ดแวร์ ที่ใช้ในการบีบอัดข้อมูล ซึ่งเป็นผลมาจากการพัฒนาเทคโนโลยีทางด้านนี้เป็นอย่างมาก เช่น RealAudio & RealVideo [1] ในประเทศไทยซึ่งมีความต้องการการใช้งานซอฟต์แวร์ประเภทนี้จำเป็นต้องนำเข้า (Import) ซอฟต์แวร์ซึ่งทำให้มีการรั่วไหลของเงินไปสู่ภายนอกเป็นจำนวนมาก

H.261 คือมาตรฐานหนึ่งที่ใช้ในการบีบอัดข้อมูลซึ่งได้เสนอไว้ใน [2] โดยมีจุดประสงค์ในการส่งผ่านข้อมูลชนิดมัลติมีเดียซึ่งมีขนาด $n \times 64$ kbps ผ่านระบบเครือข่ายโทรศัพท์ ISDN เช่น การใช้งานในระบบโทรศัพท์วีดิทัศน์ (Videophone) ต่อมาในปี 1996 ได้มีการปรับปรุงพัฒนาเทคโนโลยีดังกล่าวให้มีประสิทธิภาพสูงขึ้น และได้เปลี่ยนชื่อใหม่เป็น H.263 [3] ซึ่งจะสามารถบีบอัดข้อมูลและส่งผ่านข้อมูลไปในช่องทางสื่อสาร ที่มีขนาดเล็กกว่า 64 kbps ได้ สำหรับข้อแตกต่างระหว่างมาตรฐาน H.261 และ H.263 สามารถดูได้จาก [4] H.263 ยังได้รับการพัฒนาอย่างต่อเนื่องจนกระทั่งในปี ค.ศ. 1998 ITU-T ก็ได้ประกาศมาตรฐานที่ใช้ในการเข้ารหัสวีดิทัศน์ที่อัตราการส่งข้อมูลต่ำมากตัวใหม่ชื่อ H.263+ ซึ่งเป็นการพัฒนาเพื่อให้เหมาะกับการใช้งานในเครือข่ายประเภทเวลาจริง (Real-time) [5, 6] และขณะนี้อยู่ในระหว่างการพัฒนาต่อไปเป็นมาตรฐาน H.263++ [7] นอกจากมาตรฐานดังกล่าวมา MPEG-4 คือ วิธีการการบีบอัดข้อมูลอีกตัวหนึ่งที่มีประสิทธิภาพสูง [8] โดยได้ถูกพัฒนามาจากพื้นฐานเดียวกับ H.263 ในปัจจุบัน MPEG-4 ยังได้ถูกพัฒนาจนสามารถนำไปใช้ในระบบเครือข่ายอินเทอร์เน็ตได้อีกด้วย [9]

โครงการวิจัยนี้จึงได้ถูกเสนอมาเพื่อทำการศึกษาพัฒนาเทคโนโลยีทางด้านการบีบอัดข้อมูลชนิดวีดิทัศน์ โดยมีจุดประสงค์หลักที่จะนำผลจากการศึกษาดังกล่าว มาพัฒนาให้อยู่ในรูปแบบของโปรแกรมต้นแบบเพื่อให้สามารถใช้งานได้จริงในทางปฏิบัติ รวมไปถึงเป็นการลดการนำเข้าโปรแกรมประเภทเดียวกันจากต่างประเทศด้วย เทคโนโลยีที่ได้พัฒนาขึ้นนี้จะทำให้การส่งผ่านข้อมูลชนิดมัลติมีเดียมีประสิทธิภาพสูงขึ้น ทั้งด้านคุณภาพของข้อมูล และอัตราการส่งข้อมูลที่ต่ำกว่า นอกจากนี้ยังเป็นการปูพื้นฐานการพัฒนาเทคโนโลยีทางด้านนี้ให้กับประเทศไทย และเทคโนโลยีทางด้านที่เกี่ยวข้องในอนาคต

1.2 เรื่องที่เคยมีผู้วิจัยทำมาก่อน

การเพิ่มประสิทธิภาพการบีบอัดข้อมูลชนิดมัลติมีเดียได้มีการนำเสนออย่างมากมายในช่วงเวลา 4-5 ปีที่ผ่านมา เช่น ใน [10] Marpe and Cycon ได้เสนอการใช้วิธีการ Wavelet-based ในกระบวนการทรานฟอร์ม (Transform) แทน วิธีการ DCT-based แบบเก่า ซึ่งจะให้คุณภาพที่ดีขึ้นประมาณ 1-2 dB PSNR ที่อัตราการส่งข้อมูลเท่ากัน วิธีการดังกล่าวสามารถนำไปใช้ในมาตรฐานต่างๆ ได้ทันที เช่น ใน H.263 หรือ MPEG4

ใน [11] Pao และ Sun ได้ใช้วิธีการที่เรียกว่า Sliding-Window rate-control เพื่อรวบรวมข้อมูลทางสถิติของคุณลักษณะ (Characteristics) ของสัญญาณภาพที่ถูกบีบอัดในอดีต มาปรับค่าพารามิเตอร์สำหรับการบีบอัดข้อมูลสัญญาณภาพในอนาคต วิธีการที่คล้ายคลึงกันได้ถูกเสนอใน [12] เช่นกัน โดยที่เวลาบีบอัดข้อมูล ผู้ใช้จะป้อนพารามิเตอร์ต่างๆ เช่น ระดับสัญญาณรบกวน ลักษณะของสัญญาณ และระดับอัตราการส่งข้อมูลที่ต้องการ ให้กับตัวเข้ารหัส ซึ่งจะช่วยให้เอาท์พุทของสัญญาณที่ได้มีขนาดที่เล็กลง และมีประสิทธิภาพที่ดีขึ้นในสภาวะแวดล้อมที่ได้กำหนดเอาไว้

[13] ได้เสนอวิธีการบีบอัดข้อมูลโดยใช้เนื้อหา (Content) ของข้อมูลเป็นตัวกำหนดคุณภาพของภาพที่ได้ เช่น โปหน้าของคนจะมีความสำคัญมากกว่าส่วนอื่นๆ ของร่างกาย จึงควรใช้จำนวนบิตในการบีบอัดที่สูงกว่าส่วนอื่นๆ ของร่างกาย วิธีการนี้อาจให้ผลลัพธ์ทางอัตราการส่งข้อมูลโดยรวมที่ไม่แตกต่างจากวิธีการมาตรฐานที่ใช้อยู่ในปัจจุบัน แต่คุณภาพของภาพที่ได้จะสูงกว่า โดยเฉพาะตรงบริเวณที่สำคัญของรูปภาพ

ในโครงการวิจัยที่ได้เสนอมานี้จะทำการศึกษาและประยุกต์ใช้ วิธีการต่างๆ เพื่อให้เข้ากับการใช้งานทางด้านการบีบอัดข้อมูลที่ต้องการอัตราการส่งข้อมูลต่ำ โดยมีวัตถุประสงค์หลักก็คือ การพัฒนาโปรแกรมต้นแบบ (Software prototype) สำหรับการส่งผ่านข้อมูลชนิดมัลติมีเดียผ่านทางเครือข่ายอินเทอร์เน็ตที่มีอัตราการส่งข้อมูลที่ต่ำกว่า วิธีการที่มีอยู่ในปัจจุบัน นั่นคือ H.263 และสามารถนำไปใช้งานได้จริงในทางปฏิบัติ

1.3 วัตถุประสงค์

- เพื่อศึกษา วิธีการการเข้ารหัสสำหรับสื่อชนิดวีดิทัศน์ ที่มีอยู่ในปัจจุบัน
- เพื่อพัฒนาระบบต้นแบบที่ใช้เข้ารหัสข้อมูลชนิดวีดิทัศน์ สำหรับการนำมาใช้งานจริง เช่น การประชุมทางไกลผ่านเครือข่ายอินเทอร์เน็ต (Video conference via Internet) รวมถึงเป็นการลดปริมาณนำเข้าโปรแกรมประเภทนี้จากต่างประเทศ
- เพื่อพัฒนาโปรแกรมต้นแบบทางด้านการเข้ารหัสข้อมูลชนิดมัลติมีเดียสำหรับประเทศไทย รวมไปถึงเป็นพื้นฐานทางด้านเทคโนโลยีที่เกี่ยวข้องต่อไปในอนาคต

1.4 ขอบเขตการวิจัย

พัฒนาโปรแกรมที่ใช้ในการเข้ารหัสข้อมูลชนิดวีดิทัศน์ที่มีอัตราการส่งข้อมูลต่ำ และมีประสิทธิภาพสูงกว่าวิธีการที่มีอยู่ในปัจจุบัน โปรแกรมดังกล่าวจะถูกนำมาใช้งานบนเครื่องคอมพิวเตอร์ส่วนบุคคลซึ่งสามารถปรับแต่งขนาดของอัตราการส่งผ่านข้อมูลจาก จำนวนเฟรมของภาพต่อ 1 วินาที จำนวนส่วนประกอบของสี จำนวน GOP (Group of Picture) ของภาพ และ ความละเอียดของภาพ เป็นต้น

ตัวโปรแกรมจะถูกนำไปใช้งานทางด้านการสื่อสารทางภาพระหว่างคอมพิวเตอร์ 2 ตัวผ่านเครือข่ายอินเทอร์เน็ต เช่น ที่คอมพิวเตอร์แต่ละตัวจะมีกล้องวีดิทัศน์ขนาดเล็กอยู่ โปรแกรมดังกล่าวจะทำหน้าที่บีบอัดข้อมูลที่มาจากกล้องวีดิทัศน์ แล้วส่งผ่านไปยังหน้าจอคอมพิวเตอร์อีกตัวหนึ่ง (Non real-time applications) โดยโปรแกรมที่จะทำการพัฒนานี้จะสามารถปรับเปลี่ยนอัตราการส่งผ่านข้อมูลให้เหมาะสมกับช่องสัญญาณสื่อสารได้ นอกจากนี้โปรแกรมยังสามารถนำมาใช้งานในการบีบอัดข้อมูลที่มาจากกล้องถ่ายภาพวีดิทัศน์ระบบดิจิทัลเพื่อเก็บรักษาไว้ในคอมพิวเตอร์ได้ด้วย

1.5 ระยะเวลาที่ทำการวิจัย 1 ปี (1 ตุลาคม 2545 - 30 กันยายน 2546)

1.6 วิธีการดำเนินการวิจัยโดยสรุป

แผนการดำเนินงานโครงการวิจัยสามารถสรุปได้ดังนี้

- 1.6.1. ศึกษาวิจัย วิธีการการเข้ารหัสข้อมูลชนิดภาพเคลื่อนไหวต่าง ๆ ที่ได้มีการนำเสนอมาในอดีตรวมไปถึงโปรแกรมต่าง ๆ ที่เกี่ยวข้อง และมีการใช้งานอยู่ในปัจจุบัน
- 1.6.2. ศึกษาวิจัย วิธีการทางการประมวลผลสัญญาณชนิดต่าง ๆ ต่อข้อมูลชนิดมัลติมีเดียรวมไปถึงมาตรฐาน การเข้ารหัสข้อมูลแบบต่าง ๆ
- 1.6.3. ออกแบบระบบที่ใช้ในการเข้ารหัสข้อมูลชนิดมัลติมีเดีย
- 1.6.4. สร้างโปรแกรมต้นแบบ
- 1.6.5. ปรับปรุงตรวจสอบ แก้ไข โปรแกรมต้นแบบที่ได้สร้างขึ้น
- 1.6.6. รวบรวม ทบทวนเอกสารทั้งหมด และจัดทำรายงานสรุปผลของโครงการรวมถึงจัดทำบทความแสดงผลงานจากการวิจัย

ตารางสรุปแผนการดำเนินงาน

กิจกรรม / ระยะเวลา 1 เดือน	1	2	3	4	5	6	7	8	9	10	11	12
กิจกรรมที่ 1.6.1												
กิจกรรมที่ 1.6.2												
กิจกรรมที่ 1.6.3												
กิจกรรมที่ 1.6.4												
กิจกรรมที่ 1.6.5												
กิจกรรมที่ 1.6.6												

1.7 สถานที่ทำการทดลองและ / หรือเก็บข้อมูล

ห้องปฏิบัติการ ระบบสื่อสารมัลติมีเดีย

ภาควิชา วิศวกรรมคอมพิวเตอร์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี

1.8 แนวความคิดที่จะนำมาใช้ในการวิจัย

โครงการนี้จะทำการศึกษาวิธีการที่ใช้ในการบีบอัดข้อมูลที่ได้มีการนำเสนอ โดยกระบวนการต่าง ๆ ภายในที่ใช้ในแต่ละขั้นตอน เช่น transformation, quantization, entropy coding จะถูกนำมาพิจารณาเพื่อเลือกเอาวิธีการที่ดีที่สุดและเหมาะสมที่สุดมาประยุกต์ใช้

ข้อมูลที่จะนำมาทดสอบในการทดลองจะใช้ข้อมูลชนิดภาพเคลื่อนไหว โดยภาพเคลื่อนไหวที่ใช้ทดสอบจะเป็นภาพเคลื่อนไหวแบบมาตรฐานที่ทั่วโลกยอมรับ เช่น "bus" และ "flowers" นอกจากนี้ข้อมูลประเภทมัลติมีเดีย เช่น ข้อความหรือ ภาพเคลื่อนไหว (Animation) จะถูกนำมาทดสอบด้วย

ในการวัดผลนั้น ข้อมูลที่ถูกบีบอัดแล้วจะถูกนำมาเปรียบเทียบกับมาตรฐานการบีบอัดข้อมูล H.263 ซึ่งเป็นวิธีการที่ใช้กันอยู่ในปัจจุบัน โดยประสิทธิภาพที่เพิ่มขึ้นของวิธีการที่ได้คิดค้นและพัฒนาจะสามารถวัดได้ 2 วิธี คือ

- ที่อัตราการส่งข้อมูลเท่าเดิม คุณภาพของภาพดีขึ้น โดยวัดจากค่า Peak Signal to Noise Ratio (PSNR)
- ที่คุณภาพของภาพเท่ากัน อัตราการส่งข้อมูลต่ำกว่า โดยวัดจากค่าจำนวนบิตของข้อมูลที่ใช้ในการส่งข่าวสาร bits per second (bps)

1.9 ประโยชน์ที่คาดว่าจะได้รับ

- นำไปสู่ทฤษฎีใหม่ ความคิดใหม่ หรือวิธีใหม่ในการแก้ปัญหา
- เป็นประโยชน์ต่อประชากรกลุ่มเป้าหมาย
- เป็นพื้นฐานการวิจัยในขั้นต่อไป

1.10 หน่วยงานที่จะนำผลการวิจัยไปใช้ประโยชน์

หน่วยงานทั้งภาครัฐและเอกชนที่มีความต้องการสื่อสารข้อมูลชนิดนี้มีเดีย รวมถึงประชาชนทั่วไป

บทที่ 2 ทฤษฎีพื้นฐาน

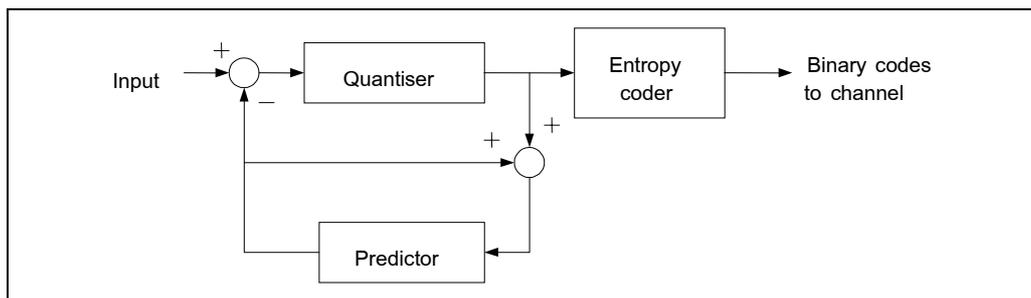
2.1 หลักการบีบอัดข้อมูลวีดิทัศน์

การบีบอัดข้อมูลวีดิทัศน์ คือการบีบอัดข้อมูลชนิดรูปภาพที่มีความต่อเนื่องกันโดยใช้ความสัมพันธ์ของข้อมูลในแต่ละเฟรมและภายในเฟรมเดียวกันเป็นหลัก จุดประสงค์ของการบีบอัดข้อมูลวีดิทัศน์นั้นก็เพื่อที่จะทำให้ข้อมูลมีขนาดเล็กที่สุด ขณะเดียวกันก็ส่งผลกระทบต่อการรับรู้ของมนุษย์ (Human perceptions) น้อยที่สุด โดยทั่วไปจะยึดหลักในการบีบอัดอยู่ 3 หลักการคือ การลดความฟุ่มเฟือยเชิงพื้นที่ (Spatial redundancy reduction) การลดความฟุ่มเฟือยเชิงเวลา (Temporal redundancy reduction) และการเข้ารหัสเอนโทรปี (Entropy coding)

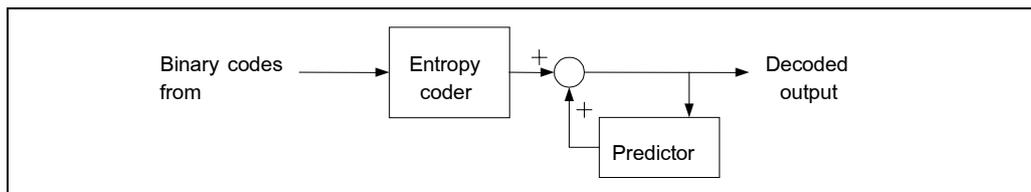
2.1.1 การลดความฟุ่มเฟือยเชิงพื้นที่ สามารถกระทำได้ 2 วิธีดังนี้

1. การเข้ารหัสแบบคาดการณ์ล่วงหน้า (Predictive coding) จะอาศัยความแตกต่างของข้อมูล ณ. เวลาปัจจุบัน เพื่อใช้คาดการณ์ลักษณะของข้อมูลในอนาคต ซึ่งสามารถทำได้ทั้งในแบบ คาดการณ์ภายในเฟรมเดียวกัน (Intraframe predictive) แบบระหว่างเฟรม (Interframe predictive) และแบบผสม (Hybrid predictive)

วิธีการที่ง่ายที่สุดที่จะนำมาใช้ในการบีบอัดข้อมูลจะอาศัยหลักการส่งข้อมูลแบบ Differential Pulse Code Modulation (DPCM) นั่นคือ การส่งข้อมูลเฉพาะส่วนที่แตกต่างกันระหว่างข้อมูลที่คาดการณ์กับข้อมูลจริงในอนาคต แผนผังแสดงหลักการทำงานของ DPCM ได้ถูกแสดงไว้ในรูปที่ 2.1



(ก)



(ข)

รูปที่ 2.1 (ก) แผนผังแสดงหลักการทำงานการเข้ารหัส DPCM (ข) การถอดรหัส DPCM

2. การแปลงข้อมูล (Transform coding) เป็นกระบวนการเปลี่ยนแปลงข้อมูลจากโดเมนเดิมไปสู่โดเมนใหม่ที่เหมาะสมต่อการลดความฟุ่มเฟือย ข้อมูลที่ถูกแปลงจะอยู่ในโดเมนความถี่ ซึ่งข้อมูลส่วนใหญ่ที่มีผลกระทบต่อสายตา มนุษย์จะอยู่ในช่วงความถี่ต่ำ ทำให้เราสามารถลดทอนข้อมูลในช่วงความถี่สูงออกไปได้ โดยทั่วไปข้อมูลที่ถูกลดทอนโดเมนนี้สามารถถูกแปลงย้อนกลับอีกครั้งมาสู่ข้อมูลในโดเมนเดิมได้ถ้าข้อมูลนั้นๆ ไม่ได้ถูกเปลี่ยนแปลงใดๆ

กระบวนการแปลงข้อมูลที่เป็นที่นิยมใช้ในข้อมูลชนิดรูปภาพ คือ การแปลงโดเมนโคซายน์แบบไม่ต่อเนื่อง (Discrete Cosine Transform, DCT) แบบ 2 มิติ (2-Dimensional) ซึ่งมีสมการดังต่อไปนี้

$$F(u,v) = \frac{C(u)}{2} \frac{C(v)}{2} \sum_{y=0}^7 \sum_{x=0}^7 f(x,y) \cos \left[\frac{(2x+1)u\pi}{16} \right] \cos \left[\frac{(2y+1)v\pi}{16} \right] \quad (2.1)$$

โดยที่ u และ v คือข้อมูลความถี่ทางแนวนอนและแนวตั้งตามลำดับ $C(u)$ และ $C(v)$ เป็นค่าคงที่ซึ่งกำหนดได้ดังนี้

$$C(v) = \frac{1}{\sqrt{2}} ; \text{ ถ้า } u = 0 \text{ และ } C(v) = 1 ; \text{ ถ้า } u > 0 \quad (2.2)$$

สมการที่ 2.1 และ 2.2 เป็นการแปลงโดเมนของจุดสี (Pixel) ภายในรูปภาพที่มีขนาด $8 \times 8 = 64$ จุด ซึ่งจะได้ผลลัพธ์เป็นค่าสัมประสิทธิ์ 64 ค่า โดยปกติข้อมูลที่ถูกแปลง ณ ตำแหน่ง (0,0) จะถูกเรียกว่า ค่าสัมประสิทธิ์ DC ส่วนข้อมูลที่ถูกแปลง ณ ตำแหน่งอื่นๆ จะถูกเรียกว่า ค่าสัมประสิทธิ์ AC การแปลงข้อมูลความถี่กลับมาในโดเมนจุดสีตามเดิมนั้น สามารถทำได้โดยการแปลงโดเมนย้อนกลับโคซายน์แบบไม่ต่อเนื่อง (Inverse Discrete Cosine Transform, IDCT) ซึ่งมีสมการดังต่อไปนี้

$$f(x,y) = \sum_{u=0}^7 \sum_{v=0}^7 \frac{C(u)}{2} \frac{C(v)}{2} F(u,v) \cos \left[\frac{(2x+1)u\pi}{16} \right] \cos \left[\frac{(2y+1)v\pi}{16} \right] \quad (2.3)$$

ในขั้นตอนการแปลงโดเมนของข้อมูลมาอยู่ในโดเมนความถี่นั้น จะยังไม่มีกระบวนการบีบอัดข้อมูลใดๆ เกิดขึ้น อย่างไรก็ตาม เนื่องจากสายตามนุษย์มีการตอบสนองต่อข้อมูลในช่วงความถี่ต่ำมากกว่าข้อมูลในช่วงความถี่สูง นั่นคือสายตามนุษย์จะสังเกตเห็นการเปลี่ยนแปลงข้อมูลในช่วงความถี่สูงได้น้อยมาก ดังนั้นเราจึงสามารถลดรายละเอียดของข้อมูลในช่วงความถี่สูงลงได้โดยใช้กระบวนการที่เรียกว่า การควอนไทเซชันค่าสัมประสิทธิ์ DCT (Quantization of DCT coefficients) หรือการกำหนดขนาดของค่าสัมประสิทธิ์ DCT

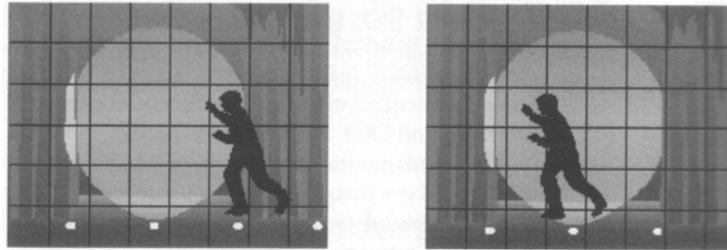
โดยทั่วไปกระบวนการควอนไทเซชันจะทำการหารค่าสัมประสิทธิ์ของข้อมูลในโดเมนเวลา $F(u,v)$ ด้วยค่าระดับควอนไทเซชัน q และจะทำการปัดค่าผลลัพธ์ที่ได้ให้เข้าใกล้จำนวนเต็มมากที่สุด ค่าที่ได้จะถูกเรียกว่า ควอนไทเซชันอินเด็ก (Quantization index) หรือ $I(u,v)$ ดังแสดงในสมการที่ 2.4 ให้สังเกตว่าการกำหนดให้ค่าระดับควอนไทเซชัน (Quantization step size) มีค่ามากๆ ซึ่งส่งผลให้ผลลัพธ์มีค่าเข้าใกล้ศูนย์

$$I(u,v) = \frac{F(u,v) \pm q}{2q} \quad (2.4)$$

ในการแปลงค่า $I(u,v)$ ย้อนกลับเพื่อหาค่า $F(u,v)$ นั้นสามารถทำได้ตามสมการที่ 2.5

$$F^q(u,v) = [2 \times I(u,v) \pm 1] \times q \quad (2.5)$$

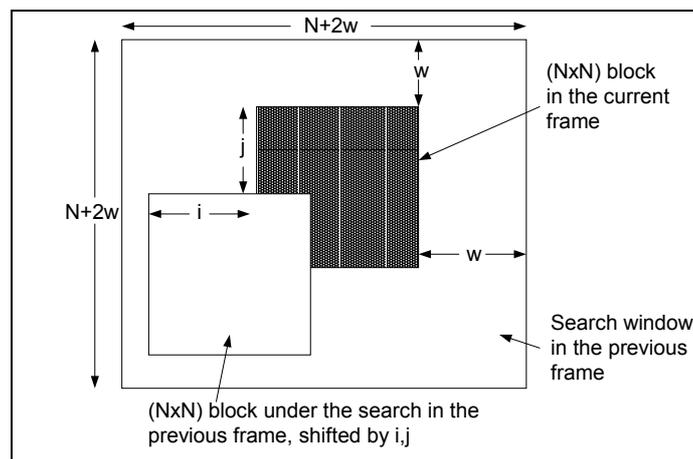
2.1.2 การลดความฟุ้งเฟ้อเชิงเวลา เป็นการลดความฟุ้งเฟ้อโดยใช้ความแตกต่างของข้อมูลชนิดรูปภาพที่มีความต่อเนื่องกัน โดยตั้งอยู่บนสมมติฐานที่ว่า รูปภาพที่มีความต่อเนื่องกันจะมีความเหมือนกันเป็นส่วนใหญ่และจะมีเพียงบางส่วนของภาพเท่านั้นที่มีการเคลื่อนไหว ตัวอย่างของรูปภาพที่มีความต่อเนื่องกันได้ถูกแสดงไว้ในรูปที่ 2.2 จะสังเกตว่า พื้นหลัง (Background) ของรูปนั้นจะเหมือนเดิมขณะที่คนในรูปจะมีการเคลื่อนไหวจากขวาไปซ้าย



รูปที่ 2.2 ตัวอย่างของรูปภาพที่มีความต่อเนื่องกัน (ภาพเคลื่อนไหว)

วิธีการที่เป็นที่นิยมใช้ในการลดความฟุ้งเฟ้อเชิงเวลานั้นก็คือ วิธีการประมาณทิศทางการเคลื่อนไหวของวัตถุในภาพ (Motion estimation) ซึ่งจะใช้พื้นฐานของการเปรียบเทียบบล็อกข้อมูล (Block matching algorithm, BMA) เป็นหลัก โดยความแตกต่างของบล็อกข้อมูลจะถูกวัดด้วย Mean squared error (MSE) หรือ Mean absolute error (MAE)

ในกระบวนการบีบอัดข้อมูลชนิดรูปภาพที่มีความต่อเนื่องกัน วิธีการเปรียบเทียบบล็อกข้อมูลจะทำการแบ่งภาพในแต่ละเฟรมออกเป็นรูปภาพย่อยขนาด $M \times N$ จุด หรือ N^2 จุด จากนั้นจะกำหนดระยะทางที่ใช้ในการหาส่วนเคลื่อนไหวเท่ากับ w จุดต่อเฟรม ข้อมูลที่อยู่ภายในบล็อกขนาด N^2 จุดนี้จะถูกเปรียบเทียบกับข้อมูลที่อยู่ภายในบล็อกขนาดเดียวกันแต่อยู่ภายในเฟรมก่อนหน้า (Previous frame) โดยระยะที่ใช้ในการหาล็อกข้อมูลเพื่อทำการเปรียบเทียบจะมีขนาดเท่ากับ $N+2w$ ดังแสดงในรูปที่ 2.3



รูปที่ 2.3 วิธีการเปรียบเทียบบล็อกข้อมูลของรูปภาพในแต่ละเฟรม

มาตรฐานที่ใช้ในการวัดความเหมือนกันของบล็อกข้อมูลในแต่ละเฟรมนั้นจะใช้ค่า MSE หรือ MAE ดังแสดงในสมการที่ 2.6 และ 2.7 ตามลำดับ

$$MSE(i, j) = \frac{1}{N^2} \sum_{m=1}^N \sum_{n=1}^N (f(m, n) - g(m + i, n + j))^2, \quad -w \leq i, j \leq w \quad (2.6)$$

$$MAE(i, j) = \frac{1}{N^2} \sum_{m=1}^N \sum_{n=1}^N |f(m, n) - g(m + i, n + j)|, \quad -w \leq i, j \leq w \quad (2.7)$$

เมื่อ $f(m, n)$ แทนบล็อกข้อมูลขนาด N^2 จุด ณ. เฟรมปัจจุบันที่ตำแหน่ง (m, n) และ $g(m + i, n + j)$ แทนบล็อกข้อมูลขนาดเดียวกัน ณ. เฟรมก่อนหน้าที่ตำแหน่ง $(m + i, n + j)$

2.1.3 การเข้ารหัสเอ็นโทรปี คือการลดความฟุ่มเฟือยด้วยวิธีการเข้ารหัสข้อมูลโดยไม่สนใจถึงความหมายของข้อมูล ข้อมูลที่เข้ารหัสจะถูกพิจารณาเป็นเพียงบิต 0 หรือ 1 เท่านั้น ในการบีบอัดข้อมูลวิธีที่ศึนั้นโดยทั่วไปจะนิยมใช้วิธีการเข้ารหัสแบบ Huffman ซึ่งจะอยู่ในรูปของรหัสที่มีความยาวไม่คงที่ (Variable Length Code, VLC)

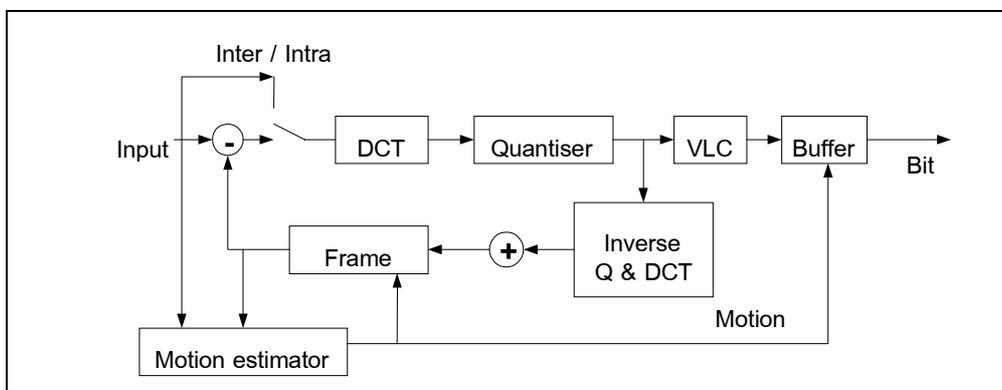
โดยหลักการของการเข้ารหัสแบบที่มีความยาวไม่คงที่นั้น ตัวเข้ารหัสจะทำการแทนค่ารหัสที่มีจำนวนบิตต่ำลงในข้อมูลที่มีความน่าจะเป็นในการเกิด (Probability of occurrence) สูง และจะทำการแทนค่ารหัสที่มีจำนวนบิตสูงลงในข้อมูลที่มีความน่าจะเป็นในการเกิดต่ำ อัตราการลดความฟุ่มเฟือยโดยวิธีการนี้สามารถหาได้จากค่าเอ็นโทรปีของตัวมันเอง ซึ่งคำนวณได้โดยใช้สมการ ที่ 2.8

$$H(x) = -\sum_{i=1}^n p_i \log_2 p_i \quad (2.8)$$

โดย ค่า p_i คือค่าความน่าจะเป็นของแต่ละสัญลักษณ์ (Symbol) ที่ถูกแทนด้วยรหัส

2.1.4 ขั้นตอนการบีบอัดข้อมูลวิธีที่ศึนั้นโดยทั่วไป

จากหลักการที่กล่าวมาข้างต้น สามารถสรุปลักษณะการบีบอัดข้อมูลวิธีที่ศึนั้นซึ่งถูกนำไปใช้ในมาตรฐานการบีบอัดข้อมูลชนิดต่างๆ เช่น H.261, H.263, MPEG-1, MPEG-2 และ MPEG-4 ได้ดังรูปที่ 2.4



รูปที่ 2.4 แผนผังการบีบอัดข้อมูลวิธีที่ศึนั้นโดยทั่วไป

2.2 มาตรฐานการเข้ารหัสวีดิทัศน์ที่มีอัตราการส่งต่ำ H.263

มาตรฐานการเข้ารหัส H.263 ถูกกำหนดขึ้นโดย International Telecommunications Union - Telecommunication Sector (ITU-T) เพื่อจุดประสงค์ในการเข้ารหัสข้อมูลวีดิทัศน์ให้มีอัตราการส่งข้อมูลต่ำ ใช้สำหรับการส่งข้อมูลวีดิทัศน์ผ่านเครือข่ายโทรศัพท์ชนิดต่างๆ เช่น Mobile Network, PSTN หรือ Narrowband ISDN

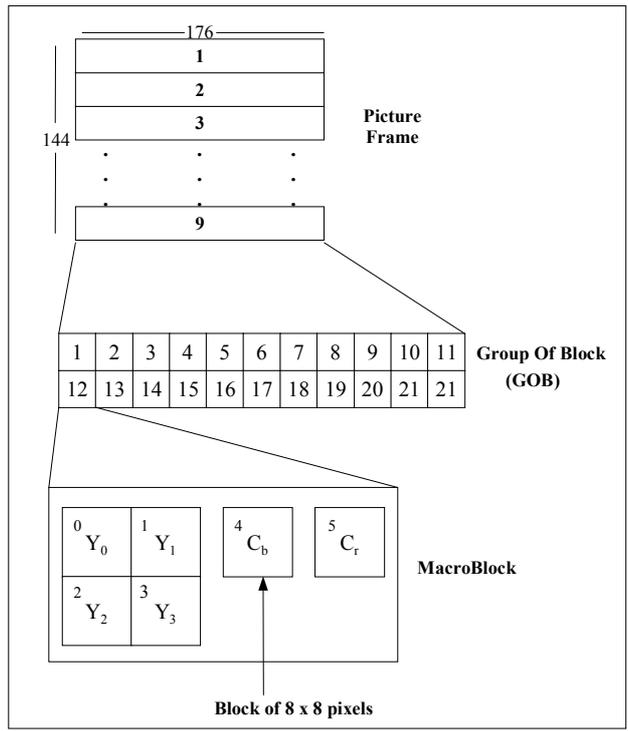
2.2.1 โครงสร้างเฟรมวีดิทัศน์ (Video Frame Structure)

H.263 รองรับมาตรฐานการเข้ารหัสวีดิทัศน์ 5 รูปแบบคือ Sub-QCIF (Sub-Common intermediate format), QCIF (Quarter Common intermediate format), CIF (Common intermediate format), 4CIF (4 Common intermediate format), 16CIF (16 Common intermediate format) ซึ่งส่วนของสัญญาณความสว่าง (Luminance) ของสัญญาณภาพจะถูกสุ่มตัวอย่าง (Sampling) ด้วยความละเอียดเท่ากับค่าความละเอียดของสัญญาณภาพ ในส่วนของสัญญาณสี (Chrominance, Cr และ Cb) จะถูกสุ่มตัวอย่างเป็นครึ่งหนึ่งของค่าความละเอียดทั้งแนวนอนและแนวตั้งของสัญญาณภาพ รายละเอียดของสัญญาณภาพทั้ง 5 รูปแบบที่ใช้ในมาตรฐานการเข้ารหัส H.263 ได้ถูกแสดงไว้ในตารางที่ 2.1

ตารางที่ 2.1 รายละเอียดของสัญญาณภาพที่ใช้ในมาตรฐานการเข้ารหัส H.263

Picture Format	Luminance		Chrominance	
	Pixel / Line	Line / Picture	Pixel / Line	Line / Picture
Sub-QCIF	128	96	64	48
QCIF	176	144	88	72
CIF	352	288	176	144
4CIF	704	576	352	288
16CIF	1408	1152	704	576

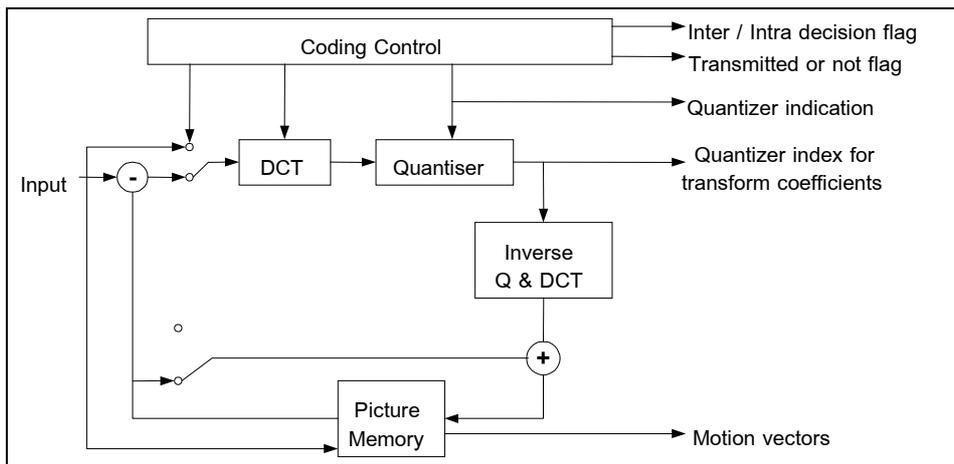
โดยทั่วไปในการเข้ารหัส ข้อมูลรูปภาพในแต่ละเฟรมจะถูกแบ่งออกเป็นกลุ่มบล็อก (Group of Block, GOB) จำนวนของ GOB จะขึ้นอยู่กับขนาดของภาพ เช่น Sub-QCIF จะมี 6 GOB, QCIF 9 และ 18 GOB สำหรับ CIF, 4CIF, 16CIF ในแต่ละ GOB จะแบ่งออกเป็นมาโครบล็อก (Macroblock, MB) ซึ่งในแต่ละ MB จะประกอบไปด้วย บล็อกข้อมูลที่เป็นสัญญาณความสว่างขนาด 8x8 จุดสี จำนวน 4 บล็อก และบล็อกข้อมูลที่เป็นสัญญาณสีขนาด 8x8 จุดสี จำนวน 2 บล็อก (Cr 1 บล็อกและ Cb 1 บล็อก) ตัวอย่างโครงสร้างเฟรมวีดิทัศน์ตามมาตรฐาน H.263 ในรูปแบบ QCIF ได้ถูกแสดงไว้ในรูปที่ 2.5



รูปที่ 2.5 ตัวอย่างโครงสร้างเฟรมวีดิทัศน์ H.263 ในรูปแบบ QCIF

2.2.2 แผนผังการเข้ารหัสวีดิทัศน์ตามมาตรฐาน H.263

รูปที่ 2.6 แสดงแผนผังการเข้ารหัสวีดิทัศน์ตามมาตรฐาน H.263 ซึ่งจะมีส่วนของการประมาณทิศทาง การเคลื่อนที่ ดังที่ได้กล่าวไว้ในส่วนของการกำจัดข้อมูลที่เหมือนกันในภาพที่มีความต่อเนื่องกัน และส่วนของการควอนไทเซชันค่าสัมประสิทธิ์ DCT



รูปที่ 2.6 แผนผังการเข้ารหัสวีดิทัศน์ตามมาตรฐาน H.263

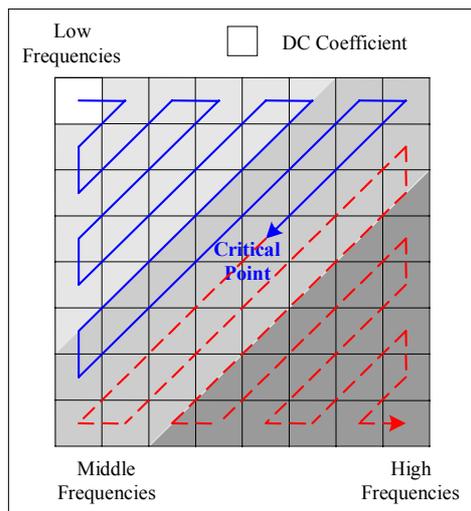
บทที่ 3 วิธีการที่นำเสนอ

เนื่องจากกระบวนการเข้ารหัสวีดิทัศน์ทุกชนิดสามารถแบ่งออกได้เป็น 2 ส่วนใหญ่ๆ คือ ส่วนของการเข้ารหัส (Encoding) และส่วนของการถอดรหัส (Decoding) การปรับปรุงประสิทธิภาพของส่วนใดส่วนหนึ่งจะส่งผลกระทบต่อรวมต่อประสิทธิภาพของการเข้ารหัสทั้งกระบวนการด้วย ในงานวิจัยนี้ได้นำเสนอวิธีการที่มุ่งเน้นการปรับปรุงประสิทธิภาพทั้งในส่วนของการเข้ารหัสและการถอดรหัส โดยในขั้นแรกได้นำหลักการของการกระจายตัวของสัมประสิทธิ์ DCT มาประยุกต์ใช้ เพื่อที่จะลดทอนขนาดของข้อมูลที่ถูกเข้ารหัสแล้วให้ที่เล็กลง โดยส่งผลกระทบต่อคุณภาพของข้อมูลภายหลังจากการถอดรหัสแล้วให้น้อยที่สุด อีกทั้งยังได้ปรับปรุงโครงสร้างข้อมูลชนิดภาพสีที่ใช้ในมาตรฐานการเข้ารหัสวีดิทัศน์ เพื่อให้มีความเหมาะสมกับงานเฉพาะด้านอีกด้วย สำหรับส่วนถอดรหัส ได้มีการปรับปรุงวิธีการประมวลผลภาพให้มีประสิทธิภาพดียิ่งขึ้น รายละเอียดของวิธีการต่างๆ จะได้แสดงในหัวข้อถัดไป

3.1 การจำกัดจำนวนสัมประสิทธิ์ DCT

จากหลักการของการเข้ารหัสข้อมูลวีดิทัศน์ตามมาตรฐาน H.263 นั้น ข้อมูลที่จะเข้ารหัสทั้งหมดจะถูกแบ่งออกเป็นบล็อกขนาด 8x8 จุดสี จากนั้นจึงนำบล็อกเหล่านั้นไปทำการแปลงโดเมนโดยใช้ DCT ที่ละบล็อก ในบล็อกหนึ่งๆ นั้น เมื่อผ่านการแปลงโดเมนและการควอนไทซ์แล้วจะได้ค่าสัมประสิทธิ์ทั้งหมด 64 ค่า รวมเรียกว่า ค่าสัมประสิทธิ์ DCT และมีการกระจายทางความถี่ในแนวทแยงมุมจากมุมบนซ้ายซึ่งเป็นความถี่ต่ำไปความถี่สูงทางมุมล่างขวา ค่าสัมประสิทธิ์ DCT นั้นยังสามารถแบ่งเป็น ค่าสัมประสิทธิ์ DC และ ค่าสัมประสิทธิ์ AC ในตำแหน่งมุมบนซ้ายคือ ค่าสัมประสิทธิ์ DC ส่วนที่เหลืออีก 63 ค่า นั้น เรียกว่า ค่าสัมประสิทธิ์ AC

เนื่องจากระบบการมองเห็นของมนุษย์ตอบสนองต่อความเปลี่ยนแปลงที่ความถี่ต่ำได้ดีกว่าความถี่สูง ดังนั้นเราจึงสามารถรับรู้ถึงความเปลี่ยนแปลงในช่วงความถี่สูงได้ยากกว่าความเปลี่ยนแปลงในช่วงความถี่ต่ำ ฉะนั้นในการนำค่าสัมประสิทธิ์ DCT ไปทำการเข้ารหัสแบบเอนโทรปีนั้น ไม่จำเป็นต้องนำค่าสัมประสิทธิ์ AC บางค่าในช่วงความถี่กลางและสูงไปทำการเข้ารหัสด้วย ซึ่งจะทำให้ข้อมูลที่นำมาเข้ารหัสนั้นมีขนาดน้อยลงและมีอัตราการส่งข้อมูลที่ต่ำลงไปด้วย

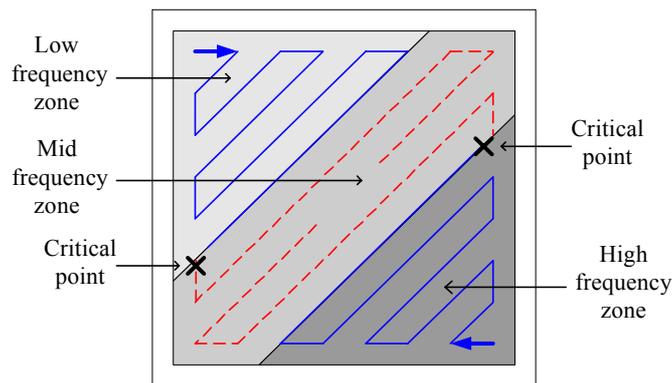


รูปที่ 3.1 สัมประสิทธิ์ DCT ที่อยู่ในช่วงเส้นทึบจะถูกนำไปเข้ารหัส

จากรูปที่ 3.1 จะเห็นว่า การสแกนค่าสัมประสิทธิ์ DCT ไปจนถึงจุดวิกฤต (Critical point) ที่กำหนดไว้ จะเป็นการกำหนดจำนวนสัมประสิทธิ์ DCT หรือระดับความลึกของสัมประสิทธิ์ DCT (Depth of DCT coefficient) ที่จะนำไปใช้ในการเข้ารหัสแบบเอนโทรปีโดยที่ไม่ทำให้คุณภาพของภาพวีดิทัศน์สูญเสียไป หรือสูญเสียน้อยที่สุดหลังจากการถอดรหัส และแทนจำนวนสัมประสิทธิ์ DCT ที่นำไปเข้ารหัสแบบเอนโทรปีด้วยสัญลักษณ์ “Z”

3.2 การสแกนแบบ 2 ทาง

ถึงแม้ว่าการไม่นำค่าสัมประสิทธิ์ DCT ในช่วงความถี่สูงบางค่ามาใช้ในการเข้ารหัสจะสามารถเพิ่มประสิทธิภาพของการเข้ารหัสข้อมูลได้ แต่ในทางทฤษฎีของการประมวลผลภาพนั้น ค่าสัมประสิทธิ์ DCT ที่อยู่ในช่วงความถี่สูงเป็นส่วนประกอบของความคมชัดของภาพ กล่าวคือความคมชัดของภาพนั้นได้ถูก DCT แปลงไปอยู่ในช่วงของความถี่สูง ดังนั้นการที่ไม่นำสัมประสิทธิ์ในช่วงความถี่สูงมาใช้ในการเข้ารหัสอาจทำให้ความคมชัดของภาพลดลง ในงานวิจัยนี้ได้นำเสนอวิธีการสแกนแบบสองทางเพื่อทำการสแกนค่าสัมประสิทธิ์ DCT จากโซนความถี่ต่ำและโซนความถี่สูง เพื่อนำไปเข้ารหัสแบบเอนโทรปี โดยเริ่มสแกนจากมุมบนซ้ายและอีกด้านหนึ่งเริ่มจากมุมล่างขวาไปตามแนวซิกแซก ซึ่งจะเห็นว่าเส้นทางการเลือกสัมประสิทธิ์ด้านความถี่สูงจะสวนทางกับแนวซิกแซกเดิม จำนวนสัมประสิทธิ์ที่เลือกจากทั้งสองโซนที่จุดวิกฤตของแต่ละโซนจะมีจำนวนเท่ากัน ส่วน Z จะมีค่าเป็นสองเท่าของจำนวนสัมประสิทธิ์ที่นับไปถึงจุดวิกฤตในแต่ละโซนดังแสดงในรูปที่ 3.2



รูปที่ 3.2 การสแกนสองทาง

3.3 การกำจัดสัญญาณสี

โดยทั่วไปแล้วสัญญาณวีดิทัศน์ในระบบแอนาล็อกจะประกอบไปด้วยสัญญาณสีแดง (Red) เขียว (Green) และฟ้า (Blue) หรือ เรียกว่า แบบจำลองสีชนิด RGB ในงานบางประเภทที่ไม่จำเป็นต้องใช้สีในการสื่อความหมาย การที่มีสัญญาณสีอยู่จึงถือเป็นความฟุ่มเฟือยในระดับหนึ่ง แต่สัญญาณสีทั้ง 3 นี้มีความสัมพันธ์ (Correlation) สูง ดังนั้นจึงไม่สามารถตัดทอนสัญญาณสีทั้งหมดได้ ทำได้เพียงบางสัญญาณเท่านั้น

ในทางทฤษฎีแล้ว สัญญาณสีฟ้าเป็นสัญญาณที่อยู่ในช่วงความถี่ที่ระบบการมองเห็นของมนุษย์นั้นไม่ไวต่อการรับรู้ ดังนั้นจึงสามารถตัดทอนสัญญาณสีฟ้าได้ แต่อย่างไรก็ตามในงานวิจัยนี้จะได้ทดสอบผลกระทบของการตัดทอนสัญญาณสีอื่นๆ ด้วย

3.4 การจำกัดส่วนประกอบสี

ถึงแม้ว่ารูปแบบของสัญญาณสีในข้อมูลวีดิทัศน์ภาพสีที่ใช้กันโดยทั่วไปในระบบแอนะล็อกนั้นนิยมใช้สัญญาณสีแดง (Red) เขียว (Green) และ ฟ้า (Blue) หรือ เรียกว่า แบบจำลองสีชนิด RGB แต่ในมาตรฐานของสีในระบบ PAL นั้นจะใช้สัญญาณสีในรูปแบบของแบบจำลองสีชนิด YUV โดยที่ Y เป็นสัญญาณที่ใช้แสดงความสว่างของภาพ ส่วน U และ V นั้นเป็นผลต่างระหว่างสีฟ้ากับค่าความสว่างของภาพ และระหว่างสีแดงกับค่าความสว่างของภาพ ตามลำดับ ซึ่งความสัมพันธ์ระหว่างแบบจำลองทั้งสองแบบเป็นไปดังสมการข้างล่างนี้

$$\begin{aligned} Y &= 0.299 R' + 0.587 G' + 0.114 B' \\ U &= -0.147 R' - 0.289 G' + 0.492 B' = 0.492 (B' - Y) \\ V &= 0.615 R' - 0.515 G' - 0.100 B' = 0.877 (R' - Y) \end{aligned} \quad (3.1)$$

โดยที่ค่า R', B' และ G' เป็นค่าของสัญญาณสีที่ผ่านการแก้ไขค่าแกมมาแล้ว (Gamma corrected) ในงานวิจัยนี้ใช้แบบจำลองสีชนิด YCbCr ตามมาตรฐาน CCIR-601 ซึ่งใช้กับระบบที่เป็นดิจิทัลโดยทั่วไป แบบจำลองสีชนิดนี้มีความใกล้เคียงกับแบบจำลองสีชนิด YUV มาก และมีความสัมพันธ์กับแบบจำลองสีชนิด RGB ตามสมการข้างล่างนี้

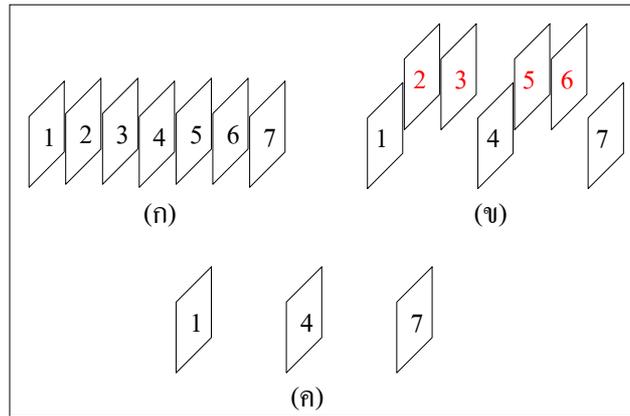
$$\begin{aligned} Y &= 0.25 R' + 0.504 G' + 0.098 B' + 16 \\ Cb &= -0.148 R' - 0.291 G' + 0.439 B' + 128 \\ Cr &= 0.439 R' - 0.368 G' - 0.071 B' + 128 \end{aligned} \quad (3.2)$$

ในการเปลี่ยนสัญญาณแอนะล็อกให้เป็นสัญญาณดิจิทัลโดยทั่วไปนั้นจะต้องทำการกรองสัญญาณ สุ่มตัวอย่างสัญญาณ และทำการกำหนดระดับขนาดของสัญญาณตามลำดับ เนื่องจากแสงสว่างมีผลต่อระบบการมองเห็นของมนุษย์มากกว่าสี กล่าวคือสายตาของคนเรานั้นสามารถรับรู้และตอบสนองต่อการเปลี่ยนแปลงของแสงสว่างได้ดีกว่าการเปลี่ยนแปลงของสี ดังนั้นในการแปลงสัญญาณของแสงและสีจากระบบแอนะล็อกเป็นดิจิทัลนั้นจึงไม่ควรใช้ความถี่ในการสุ่มตัวอย่างสัญญาณที่เท่ากัน โดยทั่วไปแล้วความถี่ของการสุ่มตัวอย่างสัญญาณแสงจะมีค่าประมาณ 2 เท่าของความถี่ในการสุ่มตัวอย่างสัญญาณสี ทั้งนี้เป็นไปตามระบบการมองเห็นของมนุษย์ดังที่ได้กล่าวไว้ข้างต้นแล้ว ฉะนั้นจึงควรเก็บรายละเอียดของแสงมากกว่าสี เพื่อให้ภาพในระบบดิจิทัลนั้นมีความใกล้เคียงกับในระบบแอนะล็อกมากที่สุดโดยอ้างอิงจากกลไกการมองเห็นดังกล่าว หลังจากแปลงข้อมูลวีดิทัศน์เป็นแบบดิจิทัลแล้ว จะต้องปรับข้อมูลภาพแต่ละภาพให้เป็นไปตามมาตรฐานภาพที่ใช้ในมาตรฐานของการเข้ารหัสวีดิทัศน์นั้นๆ ในงานวิจัยนี้ได้อ้างอิงถึงมาตรฐาน H.263+ ซึ่งเป็นมาตรฐานสำหรับการเข้ารหัสวีดิทัศน์ที่อัตราการส่งข้อมูลต่ำ ในมาตรฐานนี้ข้อมูลค่าของแสงสว่างหรือค่า Y และค่าความแตกต่างของสี Cb และ Cr จะถูกแบ่งออกเป็นบล็อกข้อมูลขนาด 8x8 จุดสี ของสีและแสงตามลำดับโดยโครงสร้างของเฟรมภาพนั้นจะเหมือนกันในรูปที่ 2.5 ที่แสดงไว้ในหัวข้อที่ 2.2.1

จากโครงสร้างของเฟรมภาพดังกล่าว หากไม่นำส่วนประกอบที่เป็นผลต่างของส่วนประกอบที่เป็นสีเทียบกับแสงสว่าง บล็อกใดบล็อกหนึ่ง หรือทั้งสองบล็อกมาใช้ในการเข้ารหัสจะสามารถทำให้ได้อัตราการส่งข้อมูลต่ำลงไปอีกซึ่งเป็นการเพิ่มประสิทธิภาพในการส่งข้อมูลให้สูงขึ้น งานวิจัยนี้ได้นำเอาหลักการข้างต้นมาทำการทดสอบ โดยทำการลดค่าความต่างของส่วนประกอบของสีเทียบกับแสงสว่าง Cb และ/หรือ Cr และหาค่าปรับแต่งที่เหมาะสมกับการใช้งานจริง

3.5 การแทรกสำเนาเฟรม

วิธีการในการลดอัตราการส่งข้อมูลที่นิยมใช้กันมากคือการลดจำนวนเฟรมภาพต่อวินาที (Frame / Sec, fps) เช่น ต้องการลดอัตราการส่งข้อมูลของวิดีโอที่ต้นฉบับที่มีอัตราเฟรม 30 เฟรมภาพต่อวินาที หากลดอัตราเฟรมภาพลงเหลือ 10 เฟรมภาพต่อวินาทีจะทำให้จำนวนข้อมูลลดลงไป 1 ใน 3 ของข้อมูลต้นฉบับ ทำให้อัตราการส่งข้อมูลลดลงไปเป็นจำนวน 1 ใน 3 ของอัตราเดิมด้วย ดังแสดงในรูปที่ 3.3 จะเห็นว่าวิธีนี้เป็นวิธีที่สามารถลดอัตราการส่งข้อมูลได้อย่างมีประสิทธิภาพ



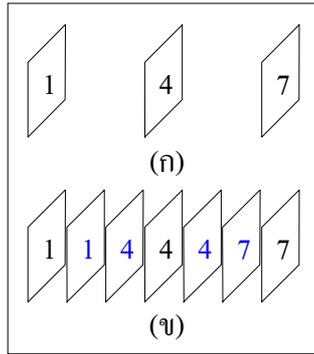
รูปที่ 3.3 (ก) ข้อมูลภาพวิดีโอต้นฉบับ (ข) เฟรมภาพที่ทำการข้ามไป (ค) เฟรมภาพที่นำไปเข้ารหัส

อย่างไรก็ตามการลดจำนวนเฟรมภาพลงมีผลทำให้ความต่อเนื่องหรือการเคลื่อนไหวซึ่งถือเป็นจุดสำคัญของข้อมูลประเภทนี้ขาดหายไปหลังจากทำการถอดรหัสข้อมูลแล้ว ยิ่งลดจำนวนเฟรมภาพจากต้นฉบับมากเท่าไรความต่อเนื่องของข้อมูลก็จะขาดหายไปมากเท่านั้น ซึ่งมีผลทำให้การสื่อสารโดยข้อมูลวิดีโอไม่ได้ประสิทธิภาพเท่าที่ควร โหมดตัวเลือกต่างๆ ที่มีอยู่ในมาตรฐาน H.263+ นั้นก็ไม่ได้มีการแก้ปัญหาในจุดนี้ ถึงแม้ว่าการหน่วงเวลาเฟรม (Frame delay) จะสามารถลดปัญหานี้ไปได้ส่วนหนึ่ง แต่ก็ไม่สามารถแก้ปัญหาได้ในระดับที่น่าพอใจ

ในงานวิจัยนี้จะได้นำเสนอวิธีการคาดเดาเฟรมที่ข้ามไป (Skipped frame prediction) ที่เรียกว่าการแทรกสำเนาเฟรม (Frame replica insertion) เพื่อนำมาใช้แก้ปัญหาความต่อเนื่องของข้อมูลที่เกิดขึ้นนี้ การหาสำเนาของเฟรมที่จะนำมาแทรกนั้นสามารถหาได้จากสมการต่อไปนี้

$$R_i(x,y) = \begin{cases} P(x,y), & i \leq d/2 \\ N(x,y), & i > d/2 \end{cases} \quad (3.3)$$

โดยที่ $R_i(x,y)$ คือสำเนาเฟรมที่ i ส่วน $P(x,y)$ และ $N(x,y)$ คือเฟรมภาพจากการถอดรหัสก่อนหน้าการแทรกเฟรมสำเนาเฟรมแรก และเฟรมภาพจากการถอดรหัสหลังจากการแทรกเฟรมสำเนาเฟรมสุดท้ายตามลำดับ d คือจำนวนของเฟรมภาพที่ทำการข้ามไม่ได้นำมาเข้ารหัสซึ่งอยู่ระหว่างเฟรมภาพ $P(x,y)$ และ $N(x,y)$ ตัวอย่างของการทำการแทรกสำเนาเฟรมได้ถูกแสดงไว้ในรูปที่ 3.4



รูปที่ 3.4 (ก) จำนวนเฟรมภาพที่มีอยู่หลังจากการถอดรหัส
 (ข) ผลลัพธ์ที่ได้หลังจากการทำกรแทรกสำเนาเฟรม

บทที่ 4 วิธีการทดลอง

ในการทดลอง ข้อมูลวีดิทัศน์ที่เป็นภาพสีจำนวนสี่ชุด ซึ่งเป็นส่วนหนึ่งของชุดข้อมูลที่ใช้เป็นมาตรฐานในการทดสอบกระบวนการเข้ารหัสข้อมูลวีดิทัศน์โดยทั่วไป ถูกนำมาใช้ในการทดสอบเพื่อหาประสิทธิภาพการเข้ารหัสวีดิทัศน์ด้วยวิธีที่นำเสนอ ซึ่งรายละเอียดของวีดิทัศน์แต่ละชุดได้แสดงไว้ในตารางที่ 4.1 รูปแบบของข้อมูลทุกชุดเป็นแบบ QCIF และมีอัตราเฟรมเป็น 30 เฟรมภาพต่อวินาที ชุดวีดิทัศน์ “Carphone” และ “Miss America” เป็นข้อมูลที่เน้นภาพใบหน้าบุคคล (Face based) เหมือนกัน แต่ต่างกันตรงที่ “Carphone” มีการเคลื่อนไหวของตัวบุคคลมากกว่า “Miss America” และมีภาพจากหลังที่เคลื่อนไหวตลอดเวลา ในขณะที่ “Miss America” ไม่มีการเคลื่อนไหวของภาพจากหลัง “Salesman” และ “Trevor” ไม่ได้เน้นรายละเอียดของใบหน้าบุคคล นอกจากนั้นแล้ว ชุดเข้ารหัสตามมาตรฐาน H.263+ ในโหมดมาตรฐานยังได้ถูกใช้ในการเข้ารหัสข้อมูลวีดิทัศน์เพื่อเปรียบเทียบกับวิธีการที่นำเสนอด้วย

ตารางที่ 4.1 รายละเอียดของชุดข้อมูลวีดิทัศน์ที่นำมาใช้ทดสอบ

Sequence Name	Frame Number	Frame Format	Video Length (Sec.)
Carphone	382	QCIF 176 x 144	12.73
Miss America	150	QCIF 176 x 144	5
Salesman	449	QCIF 176 x 144	14.97
Trevor	150	QCIF 176 x 144	5

การประเมินประสิทธิภาพของวิธีการที่นำเสนอจะวัดจากอัตราการส่งข้อมูลซึ่งหาได้จากการนำขนาดของข้อมูลที่ผ่านมาผ่านการเข้ารหัสแล้ว (kbits) มาหารด้วยความยาวของข้อมูลนั้นๆ (Second) นอกจากนี้คุณภาพของข้อมูลที่ได้จะถูกประเมินโดยใช้ค่า Peak Signal to Noise Ratio หรือ PSNR ซึ่งสามารถคำนวณหาได้จากสมการต่อไปนี้

$$PSNR = 10 \log_{10} \left[\frac{255^2}{\frac{1}{N} \sum (P_{ref}(x,y) - P_{prc}(x,y))^2} \right] \quad (4.1)$$

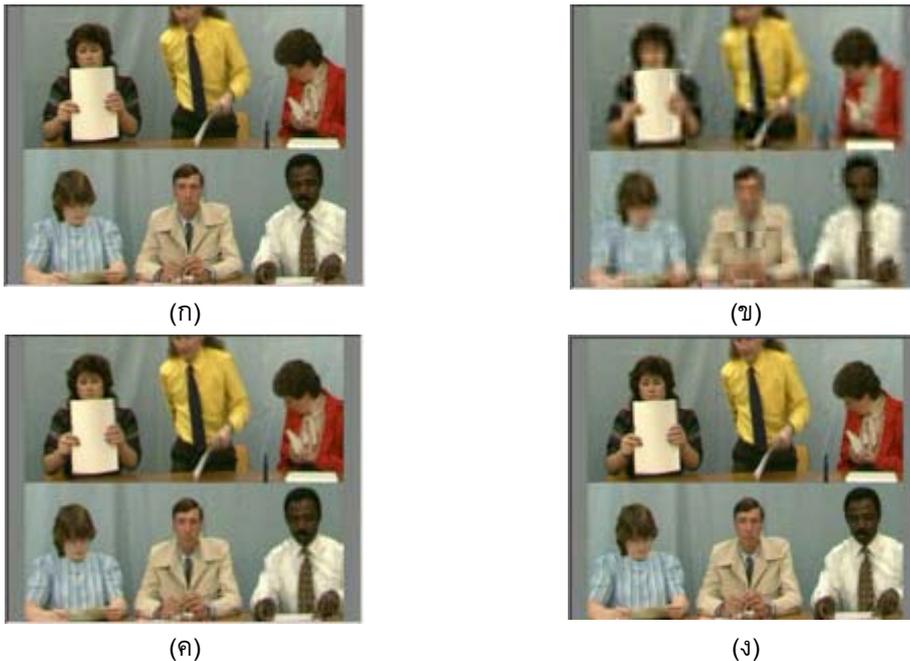
เมื่อ N คือจำนวนจุดสีทั้งหมดที่มีในรูปภาพ ส่วน $P_{ref}(x,y)$ และ $P_{prc}(x,y)$ คือค่าตัวเลขของจุดสีในภาพอ้างอิง และภาพที่ผ่านกระบวนการเข้ารหัสมาแล้วตามลำดับ ผลรวมของ PSNR จากทุกๆ เฟรมภาพจะถูกนำมารวมกัน แล้วหารด้วยค่าของจำนวนเฟรมภาพเพื่อหาผลลัพธ์ออกมาเป็นค่า PSNR เฉลี่ยของเฟรมภาพทั้งหมด จากนั้นก็จะทำการเปรียบเทียบประสิทธิภาพจากผลการทดลองของวิธีการที่นำเสนอโดยการเปรียบเทียบ PSNR กับภาพต้นฉบับ ถ้าค่า PSNR มีค่ามากแสดงว่ากระบวนการเข้ารหัสและถอดรหัสภาพสามารถถอดภาพออกมาได้เหมือนภาพต้นฉบับมาก

การตรวจสอบคุณภาพนั้นสามารถทำได้โดยใช้สายตาของคนเป็นเครื่องมือในการตรวจสอบก็ได้ ซึ่งเรียกวิธีการนี้ว่า Subjective Test ค่าที่ได้จะแปรผันตามระบบการมองเห็นของมนุษย์ ถือว่าเป็นการทดสอบที่ได้รับการยอมรับมากที่สุดเพราะเกี่ยวข้องกับความรู้สึกพอใจของมนุษย์เรา แต่เป็นการทดสอบที่ไม่ค่อยสะดวกนักเพราะต้องใช้เวลามากในการทดสอบและอาจมีความโน้มเอียง (Bias) เกิดขึ้นได้ อย่างไรก็ตามในการทดลองการจำกัดค่าสัญญาณสีและส่วนประกอบสีนั้น สิ่งที่เปลี่ยนแปลงมีเพียงแค่ส่วนประกอบที่เป็นสีเท่านั้น ดังนั้นในการตรวจสอบคุณภาพโดยวิธี Subjective Test นั้นจะใช้การประเมินจากสายตาคนทั่วไปว่าคุณภาพของภาพที่ได้มีการลดทอนลงอย่างเห็นได้ชัดหรือไม่

บทที่ 5 ผลการทดลองและข้อวิจารณ์

5.1 การจำกัดจำนวนสัมประสิทธิ์ DCT

รูปที่ 5.1 แสดงตัวอย่างผลลัพธ์ที่ได้ของเฟรมภาพที่ผ่านการเข้ารหัสโดยวิธีจำกัดจำนวนสัมประสิทธิ์ DCT โดยนำภาพมาจากชุดข้อมูล "Trevor" เฟรมภาพที่ 25 จากรูปที่ 5.1 (ก) (ค) และ (ง) พบว่าคุณภาพของภาพที่ได้นั้นเมื่อมองด้วยตาเปล่าแล้ว ไม่สามารถบ่งบอกถึงความแตกต่างอย่างชัดเจนได้ง่าย โดยเฉพาะอย่างยิ่งในขณะที่มีการเล่นกลับ (Play Back) ที่อัตราเฟรมภาพ 30 เฟรมภาพต่อวินาทีตามข้อมูลต้นฉบับ



รูปที่ 5.1 (ก) ภาพต้นฉบับและผลที่ได้จากการจำกัดจำนวนสัมประสิทธิ์ DCT ที่ (ข) 5 ค่า ($z = 5$)
(ค) 35 ค่า ($z=35$) เทียบกับ (ง) H.263+

ตารางที่ 5.1 ค่าเฉลี่ย PSNR และอัตราการส่งข้อมูลที่ี้ได้จากการจำกัดจำนวนสัมประสิทธิ์ DCT

Z	Bit-rate (kbps)	% Diff.	Avg. PSNR (dB)	% Diff.
5	140.98	-73.34	33.67	-10.57
35	418.18	-20.92	36.88	-2.05
64	528.80	0	37.65	0

ในทางตรงข้าม เราสามารถแยกแยะความแตกต่างของภาพที่แสดงในรูปที่ 5.1 (ข) เนื่องจากมีจำนวนสัมประสิทธิ์ DCT ในช่วงความถี่ต่ำเพียง 5 ตัว ทำให้เราสามารถสังเกตเห็นถึงความเปลี่ยนแปลงที่เกิดขึ้นได้โดยง่าย ในทางทฤษฎีแล้วการขาดค่าสัมประสิทธิ์ DCT ในช่วงความถี่สูงทำให้ภาพไม่คมชัด แต่เมื่อพิจารณาค่า PSNR ดังแสดงในตารางที่ 5.1 ค่าความแตกต่างของ PSNR ที่จำนวนของสัมประสิทธิ์ DCT เป็น 64 และ 35 นั้นมีค่าน้อยมากเมื่อเทียบกับค่าความแตกต่างของ PSNR ที่จำนวนของสัมประสิทธิ์ DCT เป็น 35 และ 5 นั่นคือมีการเปลี่ยนแปลงในช่วงความถี่

ต่ำมากกว่าช่วงความถี่สูง ประกอบกับสายตาของเรานั้นตอบสนองต่อการเปลี่ยนแปลงในช่วงความถี่ต่ำได้ดีกว่าการเปลี่ยนแปลงในช่วงความถี่สูง (ดูรูปที่ 5.2 ประกอบ) จึงทำให้เราเห็นความแตกต่างของการใช้สัมประสิทธิ์ DCT จำนวน 5 ค่า มากกว่าการใช้สัมประสิทธิ์ DCT จำนวน 35 ค่า ผลลัพธ์นี้แสดงให้เห็นว่าการใช้จำนวนของสัมประสิทธิ์เพียงประมาณ 55 % ของจำนวนสัมประสิทธิ์ DCT ทั้งหมดก็สามารถรักษาคุณภาพของภาพได้ในระดับที่ยอมรับได้ ทำให้เราไม่จำเป็นต้องใช้ค่าสัมประสิทธิ์ทั้งหมดในการเข้ารหัสข้อมูล



5.2 การสแกนแบบสองทาง

จากผลการทดลองในตารางที่ 5.2 แสดงให้เห็นอัตราการส่งข้อมูลและค่า PSNR แปรผันโดยตรงกับจำนวนสัมประสิทธิ์ DCT นั่นคือ หากเข้ารหัสด้วยจำนวนสัมประสิทธิ์ DCT มากเท่าไรค่า PSNR ซึ่งแสดงถึงคุณภาพของวิดีโอที่ดีขึ้นหลังจากการถอดรหัสแล้วจะสูงขึ้นเท่านั้น และอัตราการส่งข้อมูลก็จะสูงขึ้นด้วย ซึ่งไม่ใช่เป้าหมายในการเข้ารหัสข้อมูล ในทางกลับกัน หากเข้ารหัสด้วยจำนวนสัมประสิทธิ์ DCT น้อยเท่าไร อัตราการส่งข้อมูลก็จะลดลงเท่านั้น แต่คุณภาพของวิดีโอที่ดีขึ้นหลังจากการถอดรหัสก็จะลดลงด้วย

ตารางที่ 5.2 ผลลัพธ์ของการสแกนแบบสองทาง

Z	Sequence Name							
	Carphone		Miss America		Salesman		Trevor	
	Bit-rate (kbps)	PSNR (dB)	Bit-rate (kbps)	PSNR (dB)	Bit-rate (kbps)	PSNR (dB)	Bit-rate (kbps)	PSNR (dB)
20	119.58	38.36	76.51	39.28	61.00	36.15	82.74	38.70
30	158.13	39.60	102.31	41.06	81.81	36.82	110.65	40.45
40	195.53	41.58	127.66	42.84	102.75	37.86	135.90	41.78
50	226.82	43.46	149.29	44.37	120.42	39.50	157.89	43.26
60	252.83	45.95	168.57	46.19	138.53	42.78	175.77	46.15
64	259.68	47.37	175.01	46.98	142.70	45.23	180.55	47.87

จากผลการทดลองที่ได้ พบว่าค่าที่เหมาะสมในการใช้งานคือจำนวนสัมประสิทธิ์ DCT ที่ใช้ในการสแกน 40 ค่า ซึ่งจะทำให้อัตราการส่งข้อมูลลดลงได้ถึง 26 % โดยประมาณ และคุณภาพของภาพที่ได้เมื่อมองด้วยตาเปล่าแล้วจะไม่สามารถสังเกตเห็นความแตกต่างได้อย่างชัดเจน



(ก)



(ข)



(ค)

รูปที่ 5.3 ผลลัพธ์จากเฟรมแรกของชุดข้อมูลวีดิทัศน์ “Miss America” (ก) เฟรมต้นฉบับ (ข) เข้ารหัสตามมาตรฐาน H.263+ (ค) การสแกนแบบสองทางโดยใช้จำนวนสัมประสิทธิ์ DCT 40 ค่า

5.3 การกำจัดสัญญาณสี

จากตารางที่ 5.3 จะสังเกตเห็นได้ว่า ค่าเฉลี่ย PSNR ของผลการทดลองที่ได้นั้นมีค่าน้อยและต่างจากค่า PSNR ของต้นฉบับมาก แต่ทว่าจากรูปที่ 5.4 เนื้อหาของภาพนั้นยังคงอยู่เหมือนต้นฉบับทุกประการ จะแตกต่างกันก็เพียงสีที่ต่างไปเท่านั้น ซึ่งในกรณีที่มีการกำจัดสัญญาณสีฟ้าไปนั้น ภาพที่ได้มีสีที่แตกต่างจากภาพต้นฉบับน้อยมาก และอยู่ในช่วงที่ยอมรับได้ นั้นแสดงให้เห็นว่า กรณีของการกำจัดสัญญาณสีฟ้า ค่า PSNR ไม่สามารถบ่งชี้ถึงคุณภาพของภาพที่ได้หลังจากการถอดรหัสได้อย่างแม่นยำ จะสังเกตได้ว่าการกำจัดสัญญาณสีฟ้านั้นนอกจากจะได้ภาพที่มีสีใกล้เคียงกับภาพต้นฉบับที่สุดแล้ว ยังมีผลทำให้อัตราการส่งข้อมูลลดลงได้ถึง 29 % โดยประมาณอีกด้วย

ตารางที่ 5.3 ผลการทดลองการกำจัดสัญญาณสีต่างๆ จาก “Miss America”

Remaining Components	Size (bytes)	Size Diff. (%)	Average PSNR (dB)
H.263+	869,212	0.00	46.93
R	267,589	-69.21	23.12
G	525,938	-39.49	21.79
B	584,353	-32.77	18.16
RG	619,023	-28.78	27.85
RB	701,119	-19.34	21.86
GB	804,820	-7.41	24.66



(ก)



(ข)



(ค)

รูปที่ 5.4 ตัวอย่างการกำจัดสัญญาณสีจาก เฟรมภาพที่ '0' ของวิดีโอทัศน์ "Miss America" โดย
(ก) มีสัญญาณสี RG (ข) มีแต่สัญญาณสี G และ (ค) เฟรมต้นฉบับ

5.4 การกำจัดส่วนประกอบที่เป็นสี

ในตารางที่ 5.4 เป็นตัวอย่างผลการทดลองการปรับค่าส่วนประกอบสี Cb และ Cr ตัวใดตัวหนึ่งให้เป็น 0 และการปรับค่าทั้งสองให้เป็น 0 โดยใช้ชุดข้อมูล "Miss America" จากชุดวิดีโอทัศน์ทดสอบตามที่ได้ระบุไว้ข้างต้น รูปที่ 5.5 เป็นภาพผลลัพธ์ที่ได้จากการทดลองของเฟรมภาพแรกของชุดข้อมูลวิดีโอทัศน์ "Claire" และ "Miss America"

ตารางที่ 5.4 ผลการทดลองการปรับค่าส่วนประกอบสีให้เป็น 0 ของข้อมูลวิดีโอทัศน์ "Miss America"

Parameters	YcrCb	Y	Y + Cr	Y+ Cb
Size (KByte)	869,212	495,006	639,410	729,738
Play back Time (Sec.)	5	5	5	5
Encoding Time (Sec.)	34.16	34.22	34.10	34.07
Improvement	0	43.05 %	26.44 %	16.05 %



(ก)



(ข)



(ค)



(ง)

รูปที่ 5.5 เฟรมภาพผลลัพธ์จากเฟรมแรกของชุดข้อมูลวีดิทัศน์ "Miss America" ที่มีค่า
(ก) Y, Cb และ Cr (ข) Y และ Cr (ค) เฉพาะ Y (ง) Y และ Cb

จากตารางที่ 5.4 แสดงให้เห็นว่าเมื่อทำการปรับค่า Cr และ Cb ให้เป็น 0 แล้วอัตราการส่งข้อมูลจะมีค่าต่ำกว่าการปรับค่า Cr หรือ Cb เพียงอย่างเดียว อย่างไรก็ตาม จากรูปภาพจะเห็นได้ชัดว่าสีของภาพผลลัพธ์ที่เกิดจากการปรับค่า Cb มีการเปลี่ยนแปลงเพียงเล็กน้อยเท่านั้นซึ่งถือว่าอยู่ในเกณฑ์ที่ยอมรับได้ ส่วนในอีกสองกรณีคือการปรับค่าทั้ง Cr และ Cb และการปรับค่า Cr เพียงอย่างเดียวจะทำให้คุณภาพของภาพลดลงไปจนถึงจุดที่ต่ำกว่าระดับที่สามารถยอมรับได้ด้วยสายตาตามมนุษย์ วิธีการที่นำเสนอนี้ได้ใช้ทดสอบกับสัญญาณภาพ "Carphone", "Saleman" และ "Trevor" โดยการปรับค่า Cr หรือ Cb เพียงอย่างเดียวก็ให้ผลลัพธ์ที่เป็นไปในทางเดียวกันกับการทดลองด้วย "Miss America" จากผลการทดลองพบว่าอัตราการส่งข้อมูลมีแนวโน้มเป็นไปตามทฤษฎีดังที่กล่าวในหัวข้อที่ 3.4 นั่นคือเปอร์เซ็นต์การลดลงของอัตราการส่งข้อมูลหลังการปรับค่า Cr และ Cb ควรจะมีค่าประมาณ 33.33 % และการกำจัด Cr หรือ Cb เพียงอย่างเดียวควรจะทำให้บิตเรตลดลงประมาณ 16.67 % ข้อสังเกตอย่างหนึ่งก็คือหลังจากการกำจัดส่วนประกอบที่เป็นสีทั้งหมดออกไปจากสัญญาณวีดิทัศน์แล้ว ภาพผลลัพธ์ที่ได้ก็ยังคงมีสีปรากฏอยู่ ทั้งนี้เนื่องจากในระบบคอมพิวเตอร์นั้น การแสดงสัญญาณสีจะใช้แบบจำลอง RGB ดังนั้นภาพที่ได้จึงเป็นการแปลงสัญญาณภาพในแบบจำลอง YCbCr เป็น แบบจำลอง RGB ตามสมการที่ 3.2 จึงทำให้เกิดการผิดเพี้ยนของสีจากที่ควรจะเป็น

5.5 การแทรกสำเนาเฟรม

ผลการทดลองการแทรกสำเนาเฟรมของ "Carphone" และ "Trevor" ได้ถูกแสดงไว้ในตารางที่ 5.5 และรูปที่ 5.6 โดยทำการเข้ารหัสที่ อัตราเฟรมภาพ 6 เฟรมต่อวินาที

ตารางที่ 5.5 ผลลัพธ์ที่ได้จากการแทรกสำเนาเฟรม

Sequence Name	Output Bit-rate (kbps)			Average PSNR (dB)		
	H.263+	Rep.	%Diff	H.263+	Rep.	%Diff
Carphone	208.81	208.81	0%	43.86	45.29	3.3%
Trevor	437.99	437.99	0%	39.06	40.51	3.7%



(ก)



(ข)



(ค)



(ง)

รูปที่ 5.6 เฟรมภาพผลลัพธ์จากเฟรมแรกของชุดข้อมูลวีดิทัศน์ต้นฉบับ (ก) "Carphone" และ (ข) "Trevor" เทียบกับข้อมูลวีดิทัศน์ภายหลังการถอดรหัส (ค) "Carphone" และ (ง) "Trevor" ตามลำดับ

วิธีการแทรกสำเนาเฟรมนั้นจะแตกต่างจากวิธีการอื่นๆที่ได้นำเสนอมาแล้ว เนื่องจากเป็นการแทรกสำเนาเฟรมในขณะที่ทำการถอดรหัส ดังนั้นจึงไม่มีผลกระทบต่อคุณภาพของเฟรมภาพแต่ละภาพและอัตราการส่งข้อมูลดังจะเห็นได้จากตารางที่ 5.5

ถึงแม้ว่าการแทรกสำเนาเฟรมจะไม่มีผลกระทบต่อคุณภาพของเฟรมภาพแต่ละเฟรม แต่วิธีการนี้มีผลทำให้ความต่อเนื่องของเฟรมภาพมีมากขึ้น มีการเคลื่อนไหวที่เป็นลำดับต่อเนื่องมากขึ้น ซึ่งส่งผลให้คุณภาพโดยรวมของวีดิทัศน์หลังการถอดรหัสนั้นมีค่าสูงขึ้นถึง 3.5% โดยเฉลี่ย

บทที่ 6 สรุปผลการวิจัย

ความต้องการในการสื่อสารด้วยข้อมูลประเภทมัลติมีเดียมีสูงขึ้นมากในปัจจุบัน ถึงแม้ว่าข้อมูลประเภท วีดิทัศน์ เพลง รูปภาพ ต่างๆ จะสามารถส่งผ่านระบบเครือข่ายได้ แต่ข้อมูลประเภทมัลติมีเดียนี้มีขนาดที่ใหญ่มาก โดยเฉพาะข้อมูลวีดิทัศน์ ซึ่งความกว้างของช่องสัญญาณที่ใช้ในระบบเครือข่ายในปัจจุบันที่ไม่สามารถรองรับขนาดที่ใหญ่โตของข้อมูลมัลติมีเดียได้ ดังนั้นจึงมีความจำเป็นที่จะต้องมีการเข้ารหัสข้อมูลเพื่อทำการลดขนาดของข้อมูลให้มีขนาดไม่ใหญ่มากเกินไป หรือ เล็กพอที่จะส่งผ่านระบบเครือข่ายได้โดยง่าย ในงานวิจัยนี้ได้นำเอามาตรฐาน H.263 ซึ่งเป็นมาตรฐานของการเข้ารหัสเพื่อให้มีอัตราการส่งข้อมูลที่ต่ำถึงต่ำมาก มาปรับปรุงให้มีความสามารถมากขึ้น โดยในงานวิจัยนี้ได้นำเสนอแนวทางการเพิ่มประสิทธิภาพของชุดเข้ารหัส H.263 ทั้งหมด 5 แนวทาง ได้แก่ การจำกัดจำนวนสัมประสิทธิ์ DCT การสแกนแบบ 2 ทาง การลดค่าสัญญาณสี การลดค่าส่วนประกอบที่เป็นสี และการแทรกสำเนาเฟรม

การจำกัดสัมประสิทธิ์ DCT และการสแกนแบบ 2 ทางนั้น เป็นการนำคุณลักษณะของการกระจายและจัดเรียงสัมประสิทธิ์ DCT และข้อจำกัดของระบบการมองเห็นของมนุษย์มาประยุกต์ใช้เพื่อให้ได้อัตราการส่งข้อมูลที่ลดลงโดยที่คุณภาพของเฟรมภาพนั้นไม่ได้ลดลงไปด้วย หรือลดลงไปแต่ไม่สามารถสังเกตเห็นได้ด้วยสายตาของมนุษย์ จากการทดลองสามารถสรุปได้ว่าเมื่อจำกัดจำนวนสัมประสิทธิ์ DCT ให้เหลือประมาณ 55 % จะสามารถลดอัตราการส่งข้อมูลไปได้ถึง 30 % โดยเฉลี่ย ในขณะที่คุณภาพของเฟรมภาพที่ได้หลังจากการถอดรหัสไม่สูญเสียไปเลย สำหรับการสแกนแบบ 2 ทางนั้นค่าที่เหมาะสมในการทำงานคือจำนวนสัมประสิทธิ์ DCT ที่ใช้ในการสแกน 40 ค่า จะทำให้อัตราการส่งข้อมูลลดลงได้ถึง 26 % โดยประมาณ แต่อย่างไรก็ตามคุณภาพของภาพที่ได้หลังจากการถอดรหัสก็ลดลงไปด้วยแต่ยังอยู่ในเกณฑ์ที่ยอมรับได้

ในบางกรณีที่เราไม่จำเป็นต้องใช้สีในการสื่อความหมายแต่ข้อมูลที่เราเป็นนั้นเป็นข้อมูลที่เป็นภาพสี เราสามารถใช้การลดค่าสัญญาณสี และการลดส่วนประกอบที่เป็นสี ในการทำให้อัตราการส่งข้อมูลลดลงได้ โดยไม่มีผลกระทบต่อเนื้อหาของภาพ ในการลดค่าสัญญาณสีนั้น การลดค่าสัญญาณสีฟ้าเป็นทางเลือกที่เหมาะสมที่สุด เนื่องจากสัญญาณสีฟ้าเป็นสัญญาณที่อยู่ในช่วงความถี่ที่ไม่มีผลกระทบต่อสายตาของมนุษย์มากนักทำให้เราสังเกตเห็นความเปลี่ยนแปลงได้น้อยมาก และทำให้อัตราการส่งข้อมูลลดลงไปได้ถึง 29 % โดยเฉลี่ย สำหรับวิธีการลดค่าส่วนประกอบที่เป็นสีนั้นได้ประยุกต์ลักษณะโครงสร้างข้อมูลภาพมาตรฐานสำหรับข้อมูลวีดิทัศน์มาใช้ในการลดค่าส่วนประกอบสี จากการทดลองสามารถสรุปได้ว่า การลดค่าส่วนประกอบทั้ง Cb และ Cr ให้เป็นศูนย์นั้น จะสามารถลดอัตราการส่งข้อมูลได้ถึง 33.3 % โดยประมาณ อย่างไรก็ตามคุณภาพที่ได้นั้นลดต่ำลงในระดับที่ไม่สามารถยอมรับได้ สำหรับการใช้งานในทางปฏิบัติการลดค่าส่วนประกอบ Cb จะให้ผลที่ดีกว่า Cr เนื่องจากความเปลี่ยนแปลงของสีที่เกิดขึ้นยังอยู่ในเกณฑ์ที่ยอมรับได้ และให้ผลลัพธ์ให้อัตราการส่งข้อมูลที่ลดลงไปประมาณ 26 %

ในกรณีที่เราต้องการอัตราการส่งข้อมูลที่ต่ำหรือต่ำมาก โดยทั่วไปแล้วการลดจำนวนเฟรมจะถูกนำมาใช้เพื่อลดอัตราการส่งข้อมูลให้ได้มาก แต่ทว่าวิธีการนี้จะทำให้ความต่อเนื่องของข้อมูลขาดหายไป อาจจะทำให้ความหมายที่เราต้องการสื่อสารเพี้ยนไปด้วย ถึงแม้ว่าจะใช้วิธีการห่วงเวลาเฟรมมาช่วยในการแก้ปัญหา แต่ก็สามารถแก้ได้เพียงในระดับหนึ่งเท่านั้น การแทรกสำเนาเฟรมที่นำเสนอในงานวิจัยนี้เป็นการพัฒนาให้วิธีการห่วงเวลาเฟรมมีประสิทธิภาพดีขึ้น โดยสามารถทำให้ข้อมูลวีดิทัศน์มีความต่อเนื่องมากขึ้น ขณะเดียวกันก็ไม่มีผลต่ออัตราการส่งข้อมูล และยังผลให้คุณภาพของวีดิทัศน์ที่ได้หลังจากการถอดรหัสเมื่อเทียบกับวีดิทัศน์ต้นฉบับแล้วดีขึ้นกว่าวิธีการห่วงเวลาเฟรม 3.5 % โดยเฉลี่ย

ทุกแนวทางที่กล่าวมาแล้วนั้นล้วนเป็นการเพิ่มเติมประสิทธิภาพของส่วนเข้ารหัสในชุดเข้ารหัส ยกเว้นการแทรกสำเนาเฟรมที่เป็นการเพิ่มประสิทธิภาพของส่วนถอดรหัส และแต่ละวิธีการที่นำเสนอเป็นวิธีการที่ไม่ยุ่งยากซับซ้อน แต่มีประสิทธิภาพ ซึ่งสามารถนำไปประยุกต์ใช้กับมาตรฐานการเข้ารหัสแบบอื่นๆ ได้โดยทันที หรือมีการแก้ไขเพียงเล็กน้อย จึงนับเป็นการสะดวกต่อการเพิ่มเติมประสิทธิภาพของการเข้ารหัส มากกว่าแนวทางที่ยุ่งยากซับซ้อนอื่นๆ

ในการนำผลการวิจัยที่ได้ไปประยุกต์ใช้งานกับข้อมูลวีดิทัศน์นั้น แนวทางในการเลือกใช้วิธีการต่างๆ ที่ได้ทำการวิจัยสามารถสรุปได้ดังต่อไปนี้

ตารางที่ 6.1 การประยุกต์ใช้งานผลการวิจัย

วิธีการ	การประยุกต์ใช้	ผลกระทบ	ผลลัพธ์ที่ดีที่สุด
1. การจำกัดจำนวนสัมประสิทธิ์ DCT	ใช้ได้กับวีดิทัศน์ทุกประเภท	ทำให้ความคมชัดของภาพลดต่ำลง	เมื่อใช้จำนวนสัมประสิทธิ์ DCT เพียง 55 % ในการเข้ารหัส จะทำให้ลดอัตราการส่งข้อมูลลงประมาณ 30 % โดยเฉลี่ยโดยคุณภาพที่ลดลงไม่สามารถสังเกตได้ด้วยตาเปล่า
2. การสแกนแบบสองทาง	ใช้ได้กับวีดิทัศน์ทุกประเภท	ทำให้ความคมชัดของภาพลดต่ำลง	เมื่อใช้จำนวนสัมประสิทธิ์ DCT เพียง 62.5 % ในการเข้ารหัส จะทำให้ลดอัตราการส่งข้อมูลลงประมาณ 26 % โดยเฉลี่ยโดยคุณภาพที่ลดลงไม่สามารถสังเกตได้ด้วยตาเปล่า
3. การกำจัดสัญญาณสี	ใช้ได้กับวีดิทัศน์ทุกประเภทที่ไม่เน้นความถูกต้องของสี	ทำให้สีของข้อมูลในวีดิทัศน์ผิดเพี้ยนไปจากต้นฉบับ	เมื่อใช้สัญญาณสีเพียง RG เท่านั้น จะทำให้ลดอัตราการส่งข้อมูลลงประมาณ 29 % โดยโดยคุณภาพของวีดิทัศน์ที่ได้มีความใกล้เคียงกับต้นฉบับ และสามารถยอมรับได้ในการนำไปใช้งานเมื่อสังเกตด้วยตาเปล่า
4. การกำจัดส่วนประกอบที่เป็นสี	ใช้ได้กับวีดิทัศน์ทุกประเภทที่ไม่เน้นความถูกต้องของสี	ทำให้สีของข้อมูลในวีดิทัศน์ผิดเพี้ยนไปจากต้นฉบับ	เมื่อใช้จำนวนส่วนประกอบของสีเพียง YCr เท่านั้น จะทำให้ลดอัตราการส่งข้อมูลลงประมาณ 26 % โดยโดยคุณภาพของวีดิทัศน์ที่ได้มีความใกล้เคียงกับต้นฉบับ และสามารถยอมรับได้ในการนำไปใช้งานเมื่อสังเกตด้วยตาเปล่า
5. การแทรกสำเนาเฟรม	ใช้ได้กับวีดิทัศน์ทุกประเภทที่ถูกเข้ารหัสโดยวิธีการข้ามเฟรม	ไม่มี	จะเพิ่มความต่อเนื่อง (Smoothness) ของวีดิทัศน์ขณะที่มีการเล่นกลับ รวมทั้งคุณภาพที่ได้โดยรวมดีขึ้นประมาณ 3.5 %

เอกสารอ้างอิง

- [1] RealVideo and RealAudio8 Recommended Encoding Setting,
- [2] <http://www.realnetworks.com/devzone/documentation/index.html>
- [3] 'Video coding for low bitrate communication', ITU-T Recommendation H.263, Mar. 1996.
- [4] Girod, B., Steinbach, E. and Farber, N., 'Comparison of the H.263 and H.261 video compression standards', SPIE Proceedings Vol. CR60, Standards and Common Interfaces for Video Information System, October 25-26, 1995, Philadelphia, USA.
- [5] ITU Telecom. Standardization Sector of ITU, "Video coding for low bitrate communication", ITU-T Recommendation H.263 version 2, Sept. 1997.
- [6] G. Cote, B. Erol, M. Gallant and F. Kossentini, "H.263+ Video Coding Standard: Complexity and Performance," DCC '98 Proceedings. Data Compression Conference, 1998, pp. 259-268.
- [7] ITU Telecom. Standardization Sector of ITU, "Draft for H.263++, Annexes U, V and W to Recommendation H.263", ITU-T Recommendation H.263, February 2002.
- [8] Kneip, J.; Schmale, B. and Moller, H., 'Applying and implementing the MPEG-4 multimedia standard', IEEE Micro., Volume: 19, Issue: 6, Nov.-Dec. 1999, Page(s): 64 -74.
- [9] Wu, D., Hou, Y.T., Zhu, W., Zhang, Y.Q. and Chao, H.J., 'MPEG4 compressed video over the Internet', Proceedings of the 1999 IEEE International Symposium on Circuits and Systems (ISCAS '99), Volume: 4, 1999, Page(s): 327 -331.
- [10] Marpe, D. and Cycon, H.L., 'Very low bit-rate video coding using wavelet-based techniques', IEEE Transactions on Circuits and Systems for Video Technology, Volume: 9, Issue: 1, Feb. 1999, Page(s): 85 -94.
- [11] Pao, I.-M. and Ming-Ting Sun, 'Encoding stored video for streaming applications', IEEE Transactions on Circuits and Systems for Video Technology, Volume: 11, Issue: 2, Feb. 2001, Page(s): 199 -209.
- [12] Sangyoun, L. and Madisetti, V.K., 'Parameter optimization of robust low-bit-rate video coders', IEEE Transactions on Circuits and Systems for Video Technology, Volume: 9, Issue: 6, Sept. 1999, Page(s): 849 -855.
- [13] Karlekar, J. and Desai, U.B., 'Content based very low bit rate video coding using wavelet transform', Proceedings. 1999 International Conference on Image Processing, 1999. ICIP 99., Volume: 2 , 1999, Page(s): 66 -70.

ภาคผนวก

ก. คู่มือการใช้งาน

ข. ผลงานที่ได้รับจากโครงการ

ภาคผนวก ก. รายละเอียดโปรแกรมต้นแบบ

โปรแกรมต้นแบบ *MCL Video Codec* ที่ได้พัฒนาขึ้นสามารถนำไปใช้กับคอมพิวเตอร์ส่วนบุคคลทั่วไปที่ใช้ระบบปฏิบัติการ Windows 98, ME, XP โดยมีขั้นตอนในการใช้งานดังต่อไปนี้

1. ขั้นตอนการเตรียมไฟล์ข้อมูลวีดิทัศน์

1.1 เตรียมโปรแกรมที่ใช้บันทึกวีดิทัศน์ (ทั้งภาพและเสียง) โดยเลือกขนาดของภาพที่ต้องการให้ตรงกับมาตรฐานภาพที่ใช้กันโดยทั่วไป 5 มาตรฐาน ซึ่งมีรายละเอียดดังนี้

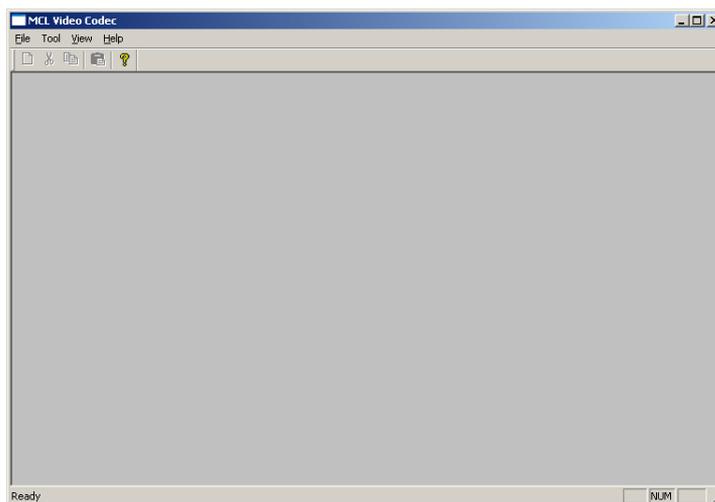
ตารางที่ ก.1 มาตรฐานภาพที่ใช้กันโดยทั่วไปในการเข้ารหัสวีดิทัศน์ที่อัตราการส่งข้อมูลต่ำ

มาตรฐานภาพ	ขนาดภาพ	
	กว้าง	ยาว
Sub-QCIF	128	96
QCIF	176	144
CIF	352	288
4CIF	704	576
16CIF	1408	1152

1.2 ทำการบันทึกภาพและเสียงจากกล้องวีดิทัศน์ โดยข้อมูลที่ทำการบันทึกจะอยู่ในรูปแบบ avi format

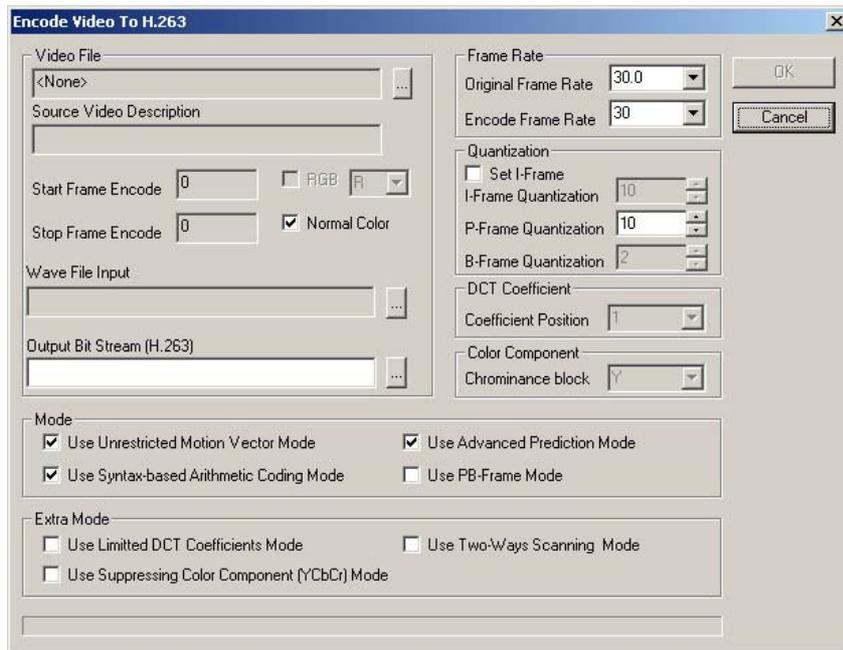
2. ขั้นตอนการเข้ารหัสวีดิทัศน์

2.1 เปิดโปรแกรม โดยการ Double Click ที่ ICON ชื่อ MCL Video Codec จะปรากฏหน้าจอตั้งรูปที่ ก.1



รูปที่ ก.1 หน้าจอหลักของโปรแกรม MCL Video Codec

2.2 ที่ Toolbar เลือก Tool → Encode Video to H.263 จะปรากฏหน้าต่างดังรูปที่ ก.2



รูปที่ ก.2 หน้าจอส่วนเข้ารหัสวีดีทัศน์

2.3 เลือกไฟล์ avi ที่ต้องการเข้ารหัส โดยทำการเลือกที่ File Type เป็น avi ก่อน จากนั้นจึงเลือกไฟล์ที่ต้องการ

2.4 ทำการเปิดไฟล์ดังกล่าว โปรแกรมจะทำการเปลี่ยน (convert) ไฟล์ดังกล่าวให้อยู่ในรูปแบบที่เหมาะสม ซึ่งประกอบด้วยไฟล์ภาพ (YUV format) และ เสียง (PCM format)

หมายเหตุ โปรแกรมจะทำการตรวจสอบขนาดของภาพโดยอัตโนมัติ ซึ่งเป็นไปตามที่ตั้งไว้ในตอนบันทึกภาพ

2.5 ใส่ชื่อไฟล์เอาท์พุทที่ต้องการลงในช่อง Output Bit Stream (H.263) ผู้ใช้สามารถเลือก Option ในการเข้ารหัสได้จากหัวข้อโหมด (Mode) (ดูรูปที่ ก.2 ประกอบ) โดยส่วนที่ได้ทำการพัฒนาขึ้นคือ Limited DCT Coefficient Mode เมื่อเลือกโหมดนี้ จะสามารถตั้งค่าสัมประสิทธิ์ที่ต้องการได้

หมายเหตุ ถ้าค่าสัมประสิทธิ์ต่ำ ค่า output bit-rate ก็จะมีค่าต่ำด้วย

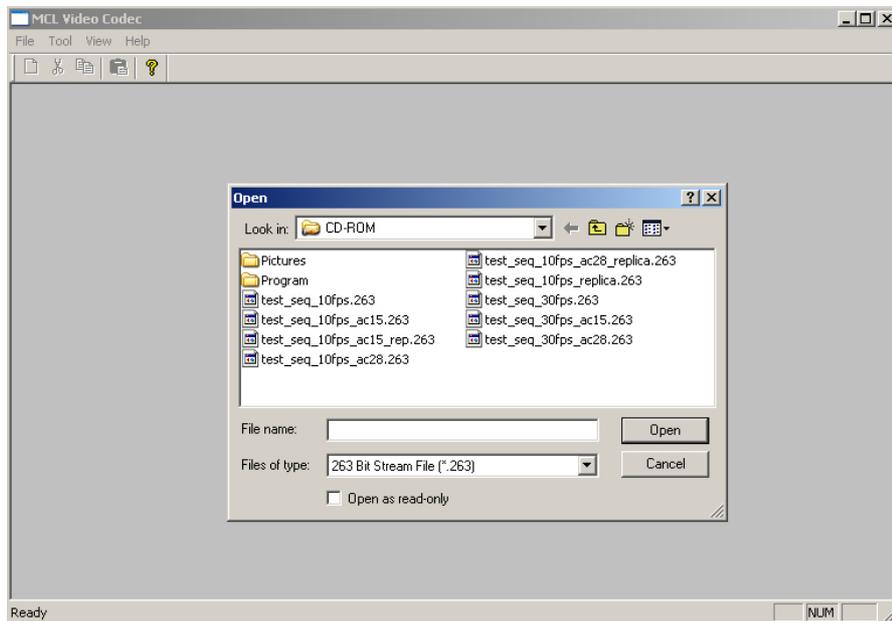
ตัวแปรที่สามารถปรับค่าได้

- Frame Rate : จำนวนเฟรมที่ต้องการใช้ในการเข้ารหัสต่อหนึ่งวินาที ค่าปกติเป็น 30 เฟรมต่อวินาที (fps)
- Start/Stop Frame : ตำแหน่งของเฟรมเริ่มต้นและเฟรมสุดท้าย ที่ใช้ในการเข้ารหัส
- P-Frame Quantization : เป็นการกำหนดระดับขนาดของสัญญาณที่จะทำการเข้ารหัส ยิ่งมีค่าสูงเท่าไร output bit-rate ยิ่งต่ำ โดยปกติถูกตั้งค่าไว้ที่ 10

2.6 เมื่อตั้งค่าต่าง ๆ ตามต้องการเสร็จแล้ว ทำการเข้ารหัสโดยกดปุ่ม OK หลังจากนั้นโปรแกรมจะสร้างไฟล์เอาท์พุทที่มีนามสกุลเป็น .263 ซึ่งประกอบไปด้วยข้อมูลภาพ และเสียง และไฟล์ตัวนี้จะถูกนำไปใช้งาน

3. ขั้นตอนการถอดรหัสวีดิทัศน์

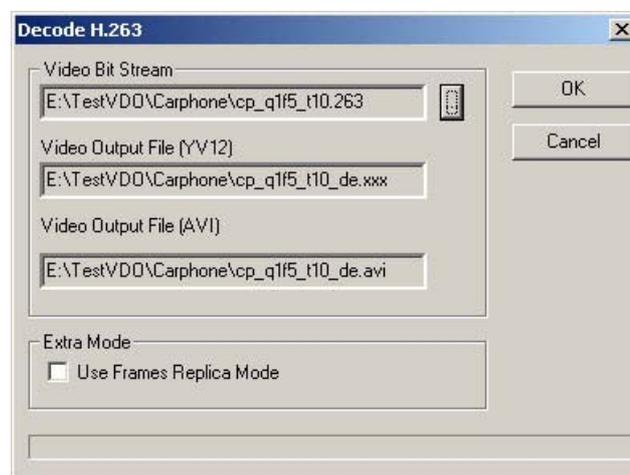
3.1 ที่ Toolbar ของหน้าจอหลักของโปรแกรม เลือก Tool → Decode H.263 Bit Stream จะปรากฏหน้าต่างดังรูปที่ ก.3



รูปที่ ก.3 เลือกไฟล์ที่ต้องการ เพื่อทำการถอดรหัส

3.2 เลือกไฟล์ที่ต้องการถอดรหัส

3.3 ในการถอดรหัส ผู้ใช้สามารถเลือกโหมดพิเศษที่ได้ทำการพัฒนาขึ้น โดยเลือกที่ Extra Mode ดังรูปที่ ก.4



รูปที่ ก.4 หน้าจอส่วนถอดรหัสวีดิทัศน์

3.4 หลังจากนั้น กดปุ่ม OK โปรแกรมจะทำการถอดรหัสและสร้างไฟล์เอาร์ทพุทให้ 2 ไฟล์ คือ ไฟล์ที่เป็นไปตามมาตรฐานภาพ 5 แบบที่ได้กำหนดไว้ และไฟล์ avi เพื่อนำไปใช้งาน

ปัญหาที่อาจเกิดขึ้นและวิธีการแก้ไข

1. ถ้าขนาดของไฟล์ avi ที่นำมาเข้ารหัส ไม่อยู่ในรูปแบบที่ได้กำหนดไว้ตามมาตรฐานภาพ 5 รูปแบบแล้ว โปรแกรม จะไม่สามารถเปลี่ยนไฟล์ให้อยู่ในรูปแบบที่เหมาะสมได้ ดังนั้น ในการบันทึกภาพวีดิทัศน์ควรจะต้องตั้งค่าความละเอียดที่เหมาะสม ตามมาตรฐานภาพ 5 รูปแบบที่กำหนดให้
2. ในกรณีที่ภาพค้างให้ตรวจสอบว่าในการบันทึกภาพได้ตั้ง frame rate ไว้ตรงกับ original frame rate ที่แสดงใน โปรแกรมหรือไม่ ถ้าไม่ต้องตั้ง frame rate ทั้งสองให้ตรงกัน

ในการทดสอบโปรแกรมได้ทำการบันทึกภาพวีดิทัศน์ในรูปแบบ ของ avi เพื่อนำมาใช้ในการทดสอบ โดยรายละเอียดของไฟล์ที่ใช้มีดังนี้

ชื่อ	test_seq1.avi
ชนิด	avi ภาพสี
ความละเอียดภาพ	QCIF 176×144
คุณภาพเสียง	PCM 44.1 KHz, 16 bit Stereo
อัตราเฟรมภาพ (frame rate)	30 fps
จำนวนเฟรม	194
เวลาเล่นกลับ	6 sec.
ขนาดของไฟล์	10 MB (10,485,760 bytes)



รูปที่ ก.5 เฟรมแรกของไฟล์ที่ใช้ในการทดสอบ

จากนั้น นำไฟล์ที่เตรียมไว้มาทำการทดสอบการเข้ารหัสและถอดรหัส โดยใช้โหมดต่างๆ ที่ได้พัฒนาขึ้น แล้วนำมา เปรียบกับการเข้ารหัสและถอดรหัสตามมาตรฐาน H.263 โดยทำการเข้ารหัสที่ 10 fps และ 30 fps ด้วยค่า P-Frame Quantization หรือ q เท่ากับ 10

ในการเข้ารหัสโดยใช้ Limited DCT Coefficients Mode นั้นได้กำหนดค่าสัมประสิทธิ์ DCT เป็น 5, 15 และ 28 ตามลำดับ ขณะทำการเข้ารหัสโปรแกรมจะแยกไฟล์จาก avi เป็นไฟล์ภาพและไฟล์เสียงตามลำดับ โดยมีรายละเอียดดังต่อไปนี้

ตารางที่ ก.2 รายละเอียดของไฟล์ที่นำมาเข้ารหัสแยกตามรูปแบบของข้อมูล

ประเภท	รูปแบบ	ขนาดของไฟล์ MB/bytes
ภาพ	QCIF 176×144	7.03 / 6,223,676
เสียง	PCM 44.1 KHz, 16 bit Stereo	1.09 / 1,151,428

เมื่อทำการถอดรหัสแล้ว นำไฟล์ที่ได้แต่ละไฟล์มาทำการหาค่า PSNR ของภาพโดยเทียบกับไฟล์ต้นฉบับ test_seq1 ได้ผลดังแสดงในตารางข้างล่างนี้

ตารางที่ ก.3 ผลลัพธ์ของการเข้ารหัสและถอดรหัสตามมาตรฐาน H.263 (Normal En / Decode)

ลักษณะการเข้ารหัส	ขนาดไฟล์ / ขนาดข้อมูลภาพ	ค่า PSNR (dB)
30 fps	1.11 MB (1,168,286 bytes) / 20 KB (16,858 bytes)	35.79
10 fps	1.10 MB (1,165,610 bytes) / 10 KB (10,210 bytes)	35.36

ตารางที่ ก.4 ผลลัพธ์ของการเข้ารหัสและถอดรหัสโดยใช้ Limited DCT Coefficients Mode

ลักษณะการเข้ารหัส	ขนาดไฟล์ / ขนาดข้อมูลภาพ	ค่า PSNR (dB)
30 fps, DCT = 5	1.11 MB (1,165,610 bytes) / 20 KB (14,182 bytes)	33.62
30 fps, DCT = 15	1.11 MB (1,167,403 bytes) / 20 KB (15,957 bytes)	34.92
30 fps, DCT = 28	1.11 MB (1,168,108 bytes) / 20 KB (16,680 bytes)	35.31
10 fps, DCT = 5	1.10 MB (1,159,833 bytes) / 10 KB (8,405 bytes)	33.40
10 fps, DCT = 15	1.10 MB (1,161,178 bytes) / 10 KB (9,750 bytes)	34.92
10 fps, DCT = 28	1.10 MB (1,161,556 bytes) / 10 KB (10,128 bytes)	35.31

ตารางที่ ก.5 ผลลัพธ์ของการเข้ารหัสและถอดรหัสโดยใช้ Frame Replica Mode

ลักษณะการเข้ารหัส	ขนาดไฟล์ / ขนาดข้อมูลภาพ	ค่า PSNR (dB)
10 fps, Rep.	1.10 MB (1,161,638 bytes) / 10 KB (16,680 bytes)	35.53

ตารางที่ ก.6 ผลลัพธ์ของการเข้ารหัสและถอดรหัสโดยใช้ Limited DCT Coefficients Mode และ Frame Replica Mode

ลักษณะการเข้ารหัส	ขนาดไฟล์ / ขนาดข้อมูลภาพ	ค่า PSNR (dB)
10 fps, DCT = 5, Rep.	1.10 MB (1,159,833 bytes) / 10 KB (8,405 bytes)	33.49
10 fps, DCT = 15, Rep.	1.10 MB (1,161,178 bytes) / 10 KB (9,750 bytes)	35.07
10 fps, DCT = 28, Rep.	1.10 MB (1,161,556 bytes) / 10 KB (10,128 bytes)	35.48

สรุปความแตกต่างระหว่างการเข้ารหัสและถอดรหัสด้วยมาตรฐาน H.263 เทียบกับโหมดที่ได้พัฒนาขึ้นเป็น เปอร์เซนต์ ณ อัตราเฟรมภาพต่อวินาที นั้นๆ

ตารางที่ ก.7 เปรียบเทียบผลลัพธ์ของ H.263 กับโหมดที่ได้พัฒนาขึ้น

ลักษณะการเข้ารหัส	Bit-rate Diff. (bps)	% Diff. Bit- rate	PSNR Diff. (dB)	% Diff. PSNR
30 fps, DCT = 5	-3568	-15.87	-2.17	-6.06
30 fps, DCT = 15	-1201.33	-5.34	-0.87	-2.43
30 fps, DCT = 28	-237.33	-1.06	-0.48	-1.34
10 fps, DCT = 5	-2406.67	-17.68	-1.96	-5.54
10 fps, DCT = 15	-613.33	-4.51	-0.44	-1.24
10 fps, DCT = 28	-109.33	-0.80	-0.05	-0.14
10 fps, Rep.	0	0	0.17	0.48
10 fps, DCT = 5, Rep.	-2406.67	-17.68	-1.87	-5.23
10 fps, DCT = 15, Rep.	-613.33	-4.51	-0.29	-0.68
10 fps, DCT = 28, Rep.	-109.33	-0.80	0.12	0.34

จากตารางที่ ก.7 สรุปได้ว่า เมื่อใช้วิธีการเข้ารหัสที่ได้พัฒนาขึ้นจะทำให้ข้อมูลที่ถูกรหัสมีขนาดลดลงประมาณ 1-18 % ขึ้นอยู่กับพารามิเตอร์ที่ผู้ใช้ได้ตั้งเอาไว้ นอกจากนี้คุณภาพของข้อมูลที่ได้จะตกลง ประมาณ 0.1-2 dB อย่างไรก็ตาม เมื่อพิจารณาจากคุณภาพของข้อมูลที่ได้โดยใช้สายตาคนทั่วไปพบว่า คุณภาพของข้อมูลที่ผ่านการเข้ารหัสสามารถยอมรับได้ในการใช้งานทางด้านมัลติมีเดียทั่วไป

หมายเหตุ เครื่องหมายลบแสดงว่าโหมดที่ได้พัฒนาขึ้นให้ค่าน้อยกว่าค่ามาตรฐาน H.263 ตัวอย่างเช่น ในการเข้ารหัสแบบ 30fps, DCT = 5 ให้ค่า output bit-rate ที่ต่ำกว่า H.263 ถึง 3,568 bps หรือประมาณ 15.87 % เป็นต้น



(ก)



(ข)



(ค)



(ง)

รูปที่ ก.6. ตัวอย่างเฟรมที่มาจากวิดีโอทดสอบ test_seq1 ที่ได้จากการเข้ารหัสสตรีมแบบต่างๆ (ก) ภาพต้นฉบับ (ข) 10 fps, q = 10 (ค) 10 fps, q = 10, DCT = 5, Rep. (ง) 10 fps, q = 10, DCT = 28, Rep.

ภาคผนวก ข. ผลงานที่ได้รับจากโครงการ

ข.1 โปรแกรมต้นแบบสำหรับการเข้ารหัสข้อมูลชนิดมัลติมีเดีย โดยใช้วิธีการที่ได้พัฒนาขึ้น ในรูปของซอฟต์แวร์

ข.2 วิทยานิพนธ์ของนักศึกษาระดับปริญญาโท 2 เล่ม

ข.3 บทความที่ได้รับการตีพิมพ์เผยแพร่ในการประชุมวิชาการระดับนานาชาติจำนวน 7 บทความ

- S. Ratanasanya, T. Amornraksa and B. Thipakorn 'Improved Performance of Very Low Bit Rate Video Coding Using Wavelet Packet Transform', Proceedings of the 17th International Technical Conference on Circuit/Systems, Computers and Communications (ITC-CSCC2002), Phuket, Thailand, July 16-19, 2002, pp. 900-903.
- S. Ratanasanya, T. Amornraksa and B. Thipakorn, 'A Smoothing Method for Reduced Frame Rate Video Sequence Using Interpolation-Based Approaches', Proceedings of the 2002 International Symposium on Communications and Information Technology (ISCIT2002), Chonburi, Thailand, October 23-25, 2002, pp. 333-336.
- P. Wongtrakulchai, S. Ratanasanya, T. Amornraksa and B. Thipakorn, 'Enhancing Performance of Video Transmission by Suppressing Color Components', Proceedings of the 2002 International Symposium on Communications and Information Technology (ISCIT2002), Chonburi, Thailand, October 23-25, 2002, pp. 337-340.
- S. Ratanasanya, S. Jiampojamarn, T. Amornraksa and B. Thipakorn, 'Video Coding Using Limited AC Coefficients and Frame Replica', Proceedings of the 2002 International Symposium on Communications and Information Technology (ISCIT2002), Chonburi, Thailand, October 23-25, 2002, pp. 388-391.
- S. Ratanasanya, T. Amornraksa and B. Thipakorn, 'Performance Comparison of Transform based Video Coding', Proceedings of the 2002 IEEE International Conference on Industrial Technology (ICIT'02), Bangkok, Thailand, December 11-14, 2002, pp. 759-763.
- S. Ratanasanya, N. Ngamwittayanon, T. Amornraksa and B. Thipakorn, 'Low Bitrate Video Coding Using Suppressed Color Components and Limited DCT Coefficients', Proceedings of the 2003 International Technical Conference on Circuit/Systems, Computers and Communications (ITC-CSCC2003), Kang-Won Do, Korea, July 7-9, 2003, Vol. I, pp. 366-369.
- S. Ratanasanya, T. Amornraksa and B. Thipakorn, 'Improved Performance of Video Transmission by Chromatics Suppressing', Proceedings of the 2003 International Technical Conference on Circuit/Systems, Computers and Communications (ITC-CSCC2003), Kang-Won Do, Korea, July 7-9, 2003, Vol. I, pp. 370-373.

ข.4 บทความที่ได้รับการตีพิมพ์เผยแพร่ในการประชุมวิชาการระดับชาติจำนวน 2 บทความ

- S. Ratanasanya, S. Jiampojamarn, T. Amornraksa and B. Thipakorn, 'Video Coding Using Limited DCT Coefficients and Frame Replica Insertion', Proceedings of the 25th Electrical Engineering Conference (EECON25), Vol. II, Songklanakarin University, Songkla, Thailand, November 21-22, 2002, pp.113-117 DS.
- N. Ngamwittayanon, S. Ratanasanya, T. Amornraksa and B. Thipakorn, 'Partial Zigzag Scanning Methods for Block-Based Video Coding', Proceedings of the 25th Electrical Engineering Conference (EECON25), Vol. II, Songklanakarin University, Songkla, Thailand, November 21-22, 2002, pp.118-122 DS.