บทที่ 1
บทนำ

1.1 แนวเหตุผลและความเป็นมา

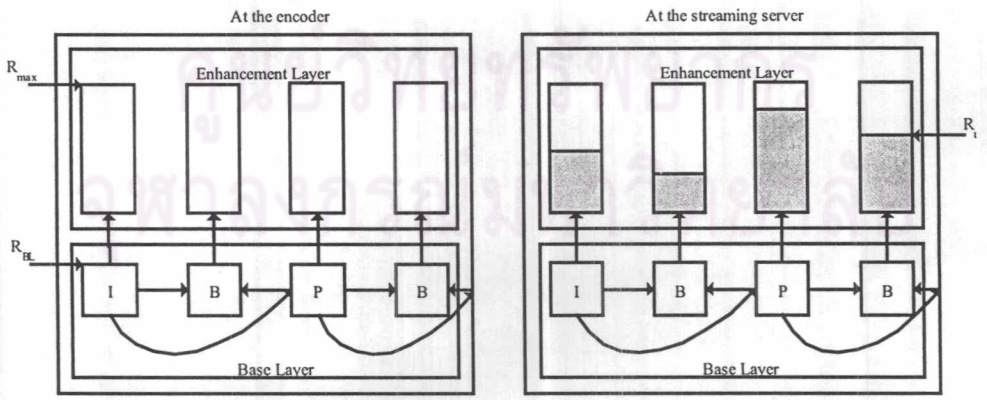
ในยุคแห่งสังคมข่าวสารเช่นปัจจุบัน การสื่อสารผ่านเครือข่ายอินเตอร์เนต ทวีความสำคัญเพิ่มขึ้นเป็นลำดับ อินเตอร์เนตช่วยให้ผู้ใช้แลกเปลี่ยนข้อมูลข่าวสารได้อย่างรวดเร็วโดยปราศจากข้อจำกัดเรื่องเวลาและระยะทาง บริการที่มีอยู่ในปัจจุบันและได้รับการตอบรับเป็นอย่างดี เช่น จดหมายอิเล็คทรอนิค (e-mail) การถ่ายโอนแฟ้มข้อมูล (ftp) การสนทนาทางเครือข่าย (chat, ICQ) และเว็บ ซึ่งเชื่อว่าในอนาคตจะมีบริการในอีกหลายลักษณะเกิดขึ้น เช่น การประชุมทางวิดีโอ (video conferencing) การศึกษาทางไกล (distance education) และทีวีผ่านเครือข่ายไอพี (IPTV) ซึ่งต้องส่งภาพวิดีโอผ่านเครือข่ายให้ได้คุณภาพดีในระดับที่น่าพอใจ เป็นต้น การส่งสัญญาณดิจิทัลวิดีโอผ่านเครือข่ายไอพีมี 2 ทางด้วยกัน คือ ยูนิแคสต์ซึ่งเป็นการส่งแบบจุดต่อจุด ในกรณีที่มีผู้รับสามคนก็ต้องส่งสามครั้ง และมัลติแคสต์เป็นการส่งแบบจุดต่อหลายจุด ถึงมีผู้รับสามคนก็ส่งเพียงครั้งเดียว โดยการส่งสัญญาณดิจิทัลวิดีโอผ่านเครือข่ายไอพีนั้น ต้องเผชิญกับ 2 ปัญหาหลักด้วยกัน [1] คือ แบนด์วิดท์ไม่คงที่ และการสูญหายของข้อมูลที่เกิดจากช่องสัญญาณมีแบนด์วิดท์ไม่คงที่ หรือ เกิดจากความคับคั่งของช่องสัญญาณหรือโครงข่าย (network congestion) ดังนั้น แนวทางการเข้ารหัสสัญญาณวิดีโอที่ต้องการในปัจจุบัน คือ สนับสนุนการเข้ารหัสที่ปรับตามสภาพของช่องสัญญาณ และการทำให้สัญญาณวิดีโอมีความทนทานต่อความผิดพลาด

MPEG-4 เป็นมาตรฐาน ISO/IEC ถูกพัฒนาโดย MPEG (Moving Picture Experts Group) ในปี ค.ศ.1998 โดยเริ่มแรกมีจุดมุ่งหมายเพื่อพัฒนามาตรฐานการเข้ารหัสให้ได้อัตราการบีบอัดข้อมูลสูงๆ ซึ่งกำหนดค่าอัตราบิตไว้ที่ต่ำกว่า 64 kbps เพื่อใช้งานในการสื่อสารข้อมูลที่อัตราบิตต่ำ หลังจากนั้นได้มีการขยายมาตรฐานเพิ่มเติมให้ครอบคลุมอัตราบิตต่ำ – สูงมาก (5kbps – 50Mbps) รวมทั้งปรับปรุงกรรมวิธีและฟังก์ชันการใช้งานมากขึ้น และยังเพิ่มความสามารถในการโต้ตอบกับผู้ใช้ เพื่อให้ควบคุมและจัดการกับภาพและเสียงได้อย่างมีประสิทธิภาพ โดยนอกจากการกำหนดการเข้ารหัสอย่างมีประสิทธิภาพแล้ว MPEG-4 ยังกำหนดรูปแบบเป็น “content-based” เพื่อให้ผู้ใช้สามารถดำเนินการกับวัตถุได้ โดยมีวัตถุประสงค์เพื่อให้สามารถทำการแสดงภาพได้อย่างยืดหยุ่น โดยเฉพาะการประยุกต์ใช้งานกับมัลติมิเดีย และด้านเกมคอมพิวเตอร์ [2][3]

ถ้าเปรียบเทียบการเข้ารหัสสัญญาณวิดีโอสำหรับซีดีรอมซึ่งเข้ารหัสตามมาตรฐาน MPEG-1 หรือการกระจายภาพทางโทรทัศน์ (TV broadcast) กับการเข้ารหัสวิดีโอสำหรับการส่งผ่านโครงข่ายอินเตอร์เนตนั้น การเข้ารหัสสำหรับอินเตอร์เนตต้องการความสามารถในการสเกลาบิลิตี้ (scalability) ที่ดีกว่า ความซับซ้อนในการคำนวณต่ำกว่า รวมทั้งมีความทนทานต่อการสูญหายของ

ข้อมูลมากกว่า กรรมวิธีการบีบอัดสัญญาณวิดีโอซึ่งเน้นการส่งสัญญาณวิดีโอผ่านช่องสัญญาณที่มีแบนด์วิดท์ไม่คงที่นั้น ส่วนใหญ่มุ่งประเด็นไปที่การเข้ารหัสแบบสเกลาบิลิตี้ซึ่งรวมอยู่ในมาตรฐาน MPEG-4 โดยจุดประสงค์หลักของการวิจัยการเข้ารหัสแบบสเกลาบิลิตี้ คือ เพื่อให้ได้ประสิทธิภาพในการบีบอัดข้อมูลที่สูงเท่าที่จะสามารถทำได้ และคาดว่าสามารถจัดการกับการเปลี่ยนแปลงของแบนด์วิดท์ได้ [4]

การเข้ารหัสและส่งสัญญาณวิดีโอด้วยวิธีสเกลาบิลิตี้ ลำดับภาพวิดีโอจะถูกเข้ารหัสเป็นสองชั้น คือ ชั้นฐาน (base layer) ซึ่งสามารถถอดรหัสและได้รับคุณภาพหยาบๆ และชั้นเอนแฮนส์เมนท์ (enhancement layer) ซึ่งถ้าสามารถถอดรหัสรวมกับชั้นฐานได้ จะได้ภาพที่มีคุณภาพที่ดีขึ้น โดยวิทยานิพนธ์นี้มุ่งประเด็นไปที่การเข้ารหัสแบบ Fine Granularity Scalability (FGS) ซึ่งเป็นส่วนหนึ่งของมาตรฐาน MPEG-4 เนื่องจากการเข้ารหัสแบบ FGS ใช้แบนด์วิดท์ของช่องสัญญาณคุ้มค่ากว่าการเข้ารหัสด้วยสเกลาบิลิตี้วิธีอื่น[1] โดยโครงสร้าง FGS ประกอบด้วย ชั้นฐาน (MPEG-4 nonscalable base layer) ที่มีอัตราการเข้ารหัส R_{BL} น้อยกว่าหรือเท่ากับแบนด์วิดท์ต่ำสุดของช่องสัญญาณ R_{min} และชั้นเอนแฮนส์เมนท์ที่เข้ารหัสด้วยอัตราการเข้ารหัส $R_{max} - R_{BL}$ โดย R_{max} คือแบนด์วิดท์สูงสุดของช่องสัญญาณ ถึงแม้จะถูกจำกัดบิตไว้เนื่องจากช่องสัญญาณมีแบนด์วิดท์ไม่พอส่งข้อมูลชั้นเอนแฮนส์เมนท์ได้ครบทั้งหมด แต่อย่างไรก็ตามสัญญาณวิดีโอที่ได้รับก็มีคุณภาพสูงกว่าที่ชั้นฐาน เพราะว่าที่เวลาจริงช่องสัญญาณจะมีแบนด์วิดท์ R_t มากกว่า R_{BL} เสมอดังแสดงในรูปที่ 1.1 ส่วนทางด้านตัวถอดรหัสนั้นก็จะถอดรหัสทั้งชั้นฐาน และบางส่วนของชั้นเอนแฮนส์เมนท์ที่ได้รับเพื่อให้ได้คุณภาพดีที่สุดเท่าที่ช่องสัญญาณจะสามารถส่งไปได้ ในขณะที่การเข้ารหัสด้วยสเกลาบิลิตี้วิธีอื่น ถ้ามีแบนด์วิดท์ไม่พอส่งข้อมูลเอนแฮนส์เมนท์ก็จะไม่มีการส่งข้อมูลเอนแฮนส์เมนท์ ทำให้คุณภาพของภาพที่ได้รับ คือ คุณภาพหยาบๆที่ได้จากชั้นฐาน



รูปที่ 1.1 ตัวอย่างการเข้ารหัสแบบ FGS สำหรับการใช้งานแบบยูนิแคสต์ [5]

การเข้ารหัสแบบ FGS นั้นมีประสิทธิภาพ มีความยืดหยุ่นเหมาะสำหรับส่งสัญญาณวิดีโอบนเครือข่ายไอพี และสามารถเลือกสเกลาบิลิตี้ที่ดีที่สุดภายหลังกระบวนการเข้ารหัส หรือขณะส่งได้ ขึ้นอยู่กับแบนด์วิดท์ของช่องสัญญาณขณะนั้น ดังนั้นการเข้ารหัสสัญญาณดิจิทัลวิดีโอแบบ

FGS จึงเป็นเรื่องน่าสนใจในการศึกษา และพัฒนากรรมวิธีการเข้ารหัสให้มีประสิทธิภาพการเข้ารหัสสูงขึ้น และเหมาะกับสภาพของช่องสัญญาณที่มีแบนด์วิดท์ไม่คงที่

งานวิจัยที่ผ่านมาในการปรับปรุงประสิทธิภาพของ FGS มี 2 วิธีหลักด้วยกัน วิธีแรก คือ ลดความซ้ำซ้อนทางเวลาภายในชั้นเอนแฮนส์เมนท์โดยการเพิ่มการชดเชยการเคลื่อนที่สำหรับเฟรม P หรือ B ที่ชั้นเอนแฮนส์เมนท์ วิธีนี้เรียกว่า Progressive FGS (PFGS) [1] หรือ Two-loop Motion-Compensation FGS (2-loop MC-FGS) [6] ซึ่งมีข้อเสียตรงที่อัตราการเข้ารหัสถูกจำกัดเนื่องจากภาพที่ชั้นฐานไม่สัมพันธ์กับอัตราบิตที่ใช้ส่ง ดังนั้นวิธีนี้จึงใช้ข้อดีในเรื่องความสัมพันธ์ทางเวลาของเฟรมต่อเนื่องกันไม่เต็มที่ และมีความซับซ้อนสูงเนื่องจากเพิ่มการชดเชยการเคลื่อนที่สำหรับเฟรม P หรือ B ที่ชั้นเอนแฮนส์เมนท์ วิธีที่สอง คือ ใช้ข้อมูลทั้งชั้นฐานและชั้นเอนแฮนส์เมนท์ในการทำนายเฟรมถัดไป เรียกว่า Single-loop Motion-Compensation FGS (1-loop MC-FGS) [6] ซึ่งทำให้ประสิทธิภาพการเข้ารหัสสูง และเนื่องจากวิธีนี้มีการชดเชยการเคลื่อนที่ที่ชั้นฐานเท่านั้นจึงมีความซับซ้อนต่ำ แต่มีข้อเสียตรงที่จะเกิดความผิดพลาดในการทำนาย (prediction drift หรือ drift error) ถ้าข้อมูลที่ใช้ในการทำนายเฟรมถัดไปไม่สามารถส่งไปได้ครบ อย่างไรก็ตาม การลดผลกระทบของความผิดพลาดในการทำนายให้ต่ำลงคือ สิ่งที่ตัวถอดรหัสต้องการ ดังนั้นในวิทยานิพนธ์นี้สนใจปรับปรุงวิธีการเข้ารหัสแบบ 1-loop MC-FGS โดยใช้ร่วมกับเทคนิคการทำนายรั่ว (leaky prediction) เพื่อเพิ่มคุณภาพสัญญาณวิดีโอให้สูงขึ้น และสามารถลดทอนความผิดพลาดที่เกิดขึ้นเมื่อมีข้อมูลบางส่วนของเฟรมก่อนหน้าสูญหาย

1.2 มาตรฐานการเข้ารหัสสัญญาณวิดีโอ

เนื่องจากสัญญาณวิดีโอมีข้อมูลขนาดใหญ่ ดังนั้นเรื่องของขนาดจึงเป็นอุปสรรคสำคัญในการส่งข้อมูลแบบวิดีโอ จึงได้มีการพัฒนาการเข้ารหัสสัญญาณวิดีโอเกิดขึ้นอย่างแพร่หลาย ในช่วง 20 ปีที่ผ่านมา มาตรฐานการเข้ารหัสวิดีโอได้รับการพัฒนาอย่างต่อเนื่อง โดยหน่วยงานที่ทำหน้าที่จัดทำมาตรฐานการเข้ารหัสสัญญาณวิดีโอมีอยู่ 2 หน่วยงานหลัก [1] ได้แก่ ITU-T (International Telecommunication Union) และ MPEG (Motion Picture Expert Group) และได้กำหนดมาตรฐานเกี่ยวกับการเข้ารหัสสัญญาณวิดีโอดังนี้

ก. มาตรฐานของ ITU-T

ITU-T เป็นมาตรฐานที่เกี่ยวข้องกับการสื่อสารโทรคมนาคมของนานาชาติ มาตรฐานการเข้ารหัสสัญญาณวิดีโอได้แก่ มาตรฐาน ITU-T ตระกูล H (H-series) จัดทำขึ้นโดยมีจุดประสงค์มุ่ง

เน้นในการใช้งานกับเครื่องโทรศัพท์ภาพ (video phone) หรือ การประชุมทางวิดีโอ (video conferencing) มาตรฐานต่าง ๆ ในตระกูล H แสดงไว้ดังตารางที่ 1.1

ตารางที่ 1.1 มาตรฐานการเข้ารหัสสัญญาณวิดีโอของ ITU-T

หมายเลขมาตรฐาน	ปีที่กำหนดใช้	วัตถุประสงค์
H.261	ค.ศ 1990	เพื่อการจัดการระบบการประชุมทางไกลผ่านระบบ ISDN (ISDN Video Conferencing)
H.262 (ร่วมกับ ISO/IEC)	ค.ศ 1995	เพื่อการส่งกระจายสัญญาณ boardcast และพัฒนาการเก็บข้อมูล มาตรฐานเหมือน MPEG 2
H.263	ค.ศ 1996	เพื่อจัดการและเข้ารหัสภาพเคลื่อนไหวผ่านระบบ PSTN และโครงข่ายไร้สาย
H.263+	ค.ศ 1998	พัฒนามาจาก H.263 โดยมีฟังก์ชันการทำงานมากขึ้น
H.263++	ค.ศ 2000	พัฒนาขึ้นจาก H.263+ และเพิ่มฟังก์ชันการทำงาน
H.26L	ค.ศ 2002	ใช้กับงานสำหรับอินเตอร์เน็ตและให้บริการสัญญาณวิดีโอ

ข. มาตรฐานของ MPEG

มาตรฐาน MPEG เป็นมาตรฐานซึ่งตั้งขึ้นโดยคณะกรรมการ MPEG ปัจจุบันเป็นมาตรฐานของ ISO (International Standard Organization) เริ่มทำงานตั้งแต่ปีคริสตศักราช 1988 โดยมีเป้าหมายเพื่อที่จะให้สามารถส่งและบันทึกสัญญาณเสียงและภาพได้เร็วและต่อเนื่อง มาตรฐาน MPEG ได้รับการยอมรับจาก ISO IEC/JTC SC29 ในปี 1991 และมีการพัฒนาเรื่อยมา ดังแสดงในตารางที่ 1.2

ตารางที่ 1.2 มาตรฐานการเข้ารหัสสัญญาณวิดีโอของ MPEG

หมายเลขมาตรฐาน	ปีที่กำหนดใช้	วัตถุประสงค์
MPEG-1	ค.ศ. 1992	เพื่อเข้ารหัสข้อมูลภาพเคลื่อนไหวสำหรับมัลติมีเดียและการเก็บข้อมูลในสื่อเก็บข้อมูล เช่น

		แผ่นซีดีรอม (Video CD)
MPEG-2	ค.ศ. 1995	พัฒนาจาก MPEG-1 เพื่อการส่งกระจายสัญญาณสำหรับโทรทัศน์ระบบดิจิทัลความละเอียดสูง (HDTV), โทรทัศน์ผ่านดาวเทียมและเคเบิ้ล รวมทั้งพัฒนาการเก็บข้อมูลให้เก็บได้มีคุณภาพสูงขึ้นและมีคุณสมบัติเพิ่มขึ้น เช่น DVD
MPEG-4	ค.ศ. 1999 (เวอร์ชัน 1) ค.ศ. 2000 (เวอร์ชัน 2)	เพื่อการส่งสัญญาณวิดีโอแบบอัตราบิตต่ำมาก (Very low bit rate) และนำไปประยุกต์ใช้งานทางอินเตอร์เน็ตแบบโต้ตอบกันได้ (Interactive) มุ่งเน้นในการบีบอัดข้อมูลและการใช้งานทางมัลติมิเดีย
MPEG-7	ค.ศ. 2001	เพื่อพัฒนามาตรฐานสู่การค้นหาและจัดเก็บข้อมูลในรูปแบบของมัลติมิเดีย
MPEG-21	ค.ศ. 2002	Multimedia Framework

1.3 แนวทางที่นำเสนอ

เนื่องจากงานวิจัยทางด้านการเข้ารหัสสัญญาณวิดีโอ มีแนวโน้มของการพัฒนาการวิจัยทางด้านการเข้ารหัสสัญญาณวิดีโอเพื่อใช้งานในการสื่อสารข้อมูลที่อัตราบิตต่ำและอัตราบิตไม่คงที่ ซึ่งมาตรฐาน MPEG-4 สามารถรองรับการใช้งานดังกล่าวได้ ดังนั้นวิทยานิพนธ์นี้จึงอยู่บนพื้นฐานการเข้ารหัสสัญญาณวิดีโอมาตรฐาน MPEG-4 โดยเข้ารหัสแบบสเกลาบิลิตี้ เพื่อรองรับช่องสัญญาณที่มีแบนด์วิดท์ไม่คงที่

การเข้ารหัสแบบสเกลาบิลิตี้ที่สนใจในวิทยานิพนธ์นี้ คือ การเข้ารหัสแบบ Fine Granularity Scalability (FGS) ซึ่งเป็นส่วนหนึ่งของมาตรฐาน MPEG-4 เนื่องจากการเข้ารหัสแบบ FGS ใช้แบนด์วิดท์ของช่องสัญญาณคุ้มค่ากว่าการเข้ารหัสด้วยสเกลาบิลิตี้วิธีอื่น[1] จุดเด่นของโครงสร้างการเข้ารหัสแบบ FGS คือ ทั้งชั้นฐาน และชั้นเอนแฮนส์เมนท์ในเฟรมปัจจุบัน ถูกทำนาย

การเคลื่อนที่จากชั้นฐานของเฟรมก่อนหน้าเท่านั้น ถึงแม้เกิดความผิดพลาด หรือเกิดการสูญหายของข้อมูลชั้นเอนแฮนส์เมนท์บางส่วน ก็ไม่มีผลกับเฟรมที่ตามมา แต่อย่างไรก็ตาม ในการทำนายจะถูกทำนายจากชั้นฐานซึ่งมีคุณภาพต่ำที่สุด ประสิทธิภาพการเข้ารหัสที่ได้รับของโครงสร้างแบบ FGS จึงไม่ดีเท่าการเข้ารหัสแบบ SNR scalability [2] ดังนั้น[6] วิธีการเข้ารหัสแบบ single-loop Motion-Compensation FGS (1-loop MC-FGS) ได้นำแนวคิดของการเข้ารหัสแบบ SNR scalability ที่ตัวเข้ารหัสและตัวถอดรหัสใช้ข้อมูลทั้งชั้นฐานและชั้นเอนแฮนส์เมนท์ในการทำนายการเคลื่อนที่เพื่อเพิ่มคุณภาพเฟรมอ้างอิงมาประยุกต์ใช้กับการเข้ารหัสแบบ FGS เพื่อปรับปรุงประสิทธิภาพการเข้ารหัส และเนื่องจากวิธีนี้มีการชดเชยการเคลื่อนที่ที่ชั้นฐานเท่านั้นจึงมีความซับซ้อนต่ำ แต่การเข้ารหัสที่ใช้ข้อมูลของชั้นเอนแฮนส์เมนท์ในการทำนายการเคลื่อนที่มีข้อเสียตรงที่จะเกิดความผิดพลาดในการทำนาย (prediction drift หรือ drift error และ error propagation) ถ้าข้อมูลที่ใช้ในการทำนายเฟรมถัดไปไม่สามารถส่งไปได้ครบ

เพื่อป้องกันและลดทอนความผิดพลาดจากการทำนาย เนื่องจากการสูญหายของข้อมูลในช่องสัญญาณที่อัตราบิตไม่คงที่ จึงเสนอเทคนิคการทำนายรั่ว (leak prediction) สำหรับวงรอบการเข้ารหัสระหว่างเฟรม (interframe loop) โดยค่าตัวประกอบการรั่ว (leakage factor : α) มีค่าตั้งแต่ 0 – 1 เทคนิคการทำนายรั่วช่วยเพิ่มความทนทานต่อความผิดพลาด (error resilience) เมื่อมีข้อมูลบางส่วนของเฟรมก่อนหน้าสูญหาย ดังนั้นในวิทยานิพนธ์นี้จึงนำเสนอการเข้ารหัสแบบ FGS โดยใช้ร่วมกับเทคนิคการทำนายรั่วที่สามารถปรับตัวได้ตามลักษณะสัญญาณวิดีโอและแบนด์วิดท์ของช่องสัญญาณ เพื่อเพิ่มคุณภาพสัญญาณวิดีโอให้สูงขึ้น

1.4 วัตถุประสงค์

เพื่อพัฒนากรรมวิธีการเข้ารหัสสัญญาณวิดีโอตามมาตรฐาน MPEG-4 เพื่อคุณภาพของสัญญาณวิดีโอที่ดีขึ้น

1.5 ขอบเขตของวิทยานิพนธ์

สามารถพัฒนาการเข้ารหัสสัญญาณวิดีโอตามมาตรฐาน MPEG-4 โดยเข้ารหัสด้วยโครงสร้างที่นำเสนอ เพื่อเพิ่มคุณภาพสัญญาณวิดีโอให้สูงขึ้นเมื่อเทียบกับการเข้ารหัสแบบ FGS เดิมและวิธี 1-loop MC-FGS และสามารถลดทอนความผิดพลาดที่เกิดขึ้นเมื่อมีข้อมูลบางส่วนสูญหาย

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

1.6 แนวทางการดำเนินงาน

1. ศึกษาความรู้พื้นฐานในการเข้ารหัสสัญญาณวิดีโอ
2. ศึกษาและวิเคราะห์การเข้ารหัสสัญญาณวิดีโอแบบไฟน์แกรนูลาริตี้สเกลาบิลิตี้
3. พัฒนาโครงสร้างเพื่อลดทอนความผิดพลาดที่นำเสนอโดยเข้ารหัสตามมาตรฐาน MPEG-4
4. จำลองช่องสัญญาณที่มีแบนด์วิดท์ไม่คงที่
5. ทดสอบโครงสร้างที่นำเสนอโดยใช้ชุดภาพทดสอบต่างๆ พร้อมทั้งวิเคราะห์ผล
6. สรุปและรวบรวมผลงานวิจัย พร้อมทั้งจัดทำรูปเล่มวิทยานิพนธ์

1.7 ประโยชน์ที่คาดว่าจะได้รับ

1. ได้ศึกษาความรู้พื้นฐานในการเข้ารหัสสัญญาณวิดีโอ และการส่งสัญญาณวิดีโอบนช่องสัญญาณที่มีแบนด์วิดท์ไม่คงที่
2. ซอฟท์แวร์อัลกอริทึมเพื่อใช้เป็นข้อมูลในการเข้ารหัส เพื่อให้ได้คุณภาพสัญญาณวิดีโอที่ดีขึ้น
3. สามารถนำไปเป็นแนวคิดในการเพิ่มประสิทธิภาพการเข้ารหัส และการลดทอนความผิดพลาด

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย