

## โครงการวิจัยย่อยลำดับที่ 9

เรื่อง การเข้ารหัสวีดิทัศน์อัตราบิตต่ำเพื่อการส่งไร้สายและงานประยุกต์เชิงวัตถุ

Project title Low Bit-Rate Video Coding for Wireless Transmission and Object-based Applications

ผู้รับผิดชอบโครงการ ผู้ช่วยศาสตราจารย์ ดร. สุภาวดี อร่ามวิทย์

### 1. ปัญหาที่ทำการวิจัยและความสำคัญของปัญหา

ในระบบการสื่อสารโทรคมนาคมยุคต่อไปขีดความสามารถรับและส่งสัญญาณหลากหลายรูปแบบ มากขึ้น ในรูปแบบของสื่อประสม ซึ่งรวมถึง ข้อมูลตัวอักษร เสียงพูด ออดิโอ ภาพนิ่ง และ วีดิทัศน์ เพื่อรองรับการใช้งานด้านต่างๆ ที่ขยายตัวอย่าง รวดเร็ว อย่างไรก็ตาม ใดก็ตามที่ช่องสัญญาณที่มีการกำหนดให้ระบบ สื่อสารโทรคมนาคม แต่ละระบบมีขนาดจำกัด ดังนั้น ปริมาณของสัญญาณจำนวนมากที่เกิดขึ้น จะทำให้เกิดความล่าช้าหรือ การประวิงเวลารวมทั้ง อาจเกิดการสูญหายของข้อมูล ซึ่งอาจก่อให้เกิดความเสียหายในการสื่อสารได้ ในสื่อประสมทั้งหมดที่กล่าวมา สัญญาณวีดิทัศน์ จัดเป็นสัญญาณที่มีปริมาณ ข้อมูลสูงที่สุด โดยเฉพาะอย่างยิ่ง ถ้าต้องการได้ภาพที่มีความละเอียดสูง

ดังนั้นงานวิจัยที่เกี่ยวข้องกับการลดรหัสสัญญาณวีดิทัศน์ จึงได้รับความสนใจ ทั้งในแวดวงการศึกษาและวงการอุตสาหกรรม รวมถึงมีการตั้งองค์กรณ์เพื่อกำหนดมาตรฐานลดรหัสสัญญาณวีดิทัศน์ เพื่อ ใช้ในการเก็บ ลงสื่อ ประเภทต่างๆ และสำหรับการส่งผ่านช่องสัญญาณ ประเภทต่างๆ แนวทางการวิจัยจะมุ่งเน้นในส่วนของการนำเสนอกรรมวิธีการเข้ารหัส ให้มีอัตราการบีบอัดข้อมูลที่สูง โดย ยัง ส า ร ถ ร ัก ษ า ร ะ ด บ ค ุ ณ ภา พ ข อ ง ส ัญ ญาณ และการปรับปรุงประสิทธิภาพในการส่งผ่านช่องสัญญาณผ่านโครงข่ายสื่อสาร ทั้งแบบมีสายและไร้สาย

ปัจจุบัน มาตรฐานการลดรหัสสัญญาณวีดิทัศน์อัตราบิตและการประวิงเวลาต่ำที่มีการใช้ ในการประยุกต์ระบบการประชุมสัมมนาทางวีดิทัศน์และการส่งสัญญาณวีดิทัศน์ผ่านเครือข่ายสื่อสาร ได้แก่ มาตรฐาน H.263 [10.1] MPEG-4 [10.2] และ มาตรฐาน H.264/MPEG-4 Part 10 [10.3] เป็นต้น มาตรฐาน MPEG-4 เป็นมาตรฐานการบีบอัดสัญญาณวีดิทัศน์เชิงวัตถุ ซึ่งให้ประสิทธิภาพการบีบอัดที่สูง เมื่อเทียบกับมาตรฐานก่อนหน้า มาตรฐาน MPEG-4 ได้ถูก

นำมาประยุกต์เชิงพาณิชย์ไม่ว่าจะเป็น ตัวบีบอัดสัญญาณวีดิทัศน์ในโทรศัพท์ไร้สายยุคที่ 3 ระบบการประชุมสัมมนาทางวีดิทัศน์ และ ระบบ Video on demand เป็นต้น สำหรับมาตรฐาน H.264/MPEG-4 Part 10 ซึ่งเป็นมาตรฐานใหม่ ในกรณีของ มาตรฐาน H.264 ซึ่งเป็นมาตรฐานใหม่ของการร่วมมือกันระหว่าง องค์กร ISO/IEC MPEG และ ITU-T VCEG (Video Coding Experts Group) ให้อัตราการบีบอัดสูง รวมทั้งคุณภาพของภาพที่ดีขึ้น ในอัตราบิตที่ต่ำกว่า ในปัจจุบัน ได้เริ่มมีการประยุกต์ใช้เชิงพาณิชย์ อย่างไรก็ตาม ประเทศไทยยังขาดผู้เชี่ยวชาญในองค์ความรู้พื้นฐานด้านนี้ ทำให้การสร้างและประยุกต์ระบบการสื่อสารสัญญาณวีดิทัศน์ผ่านเครือข่ายสื่อสารต้องมีการพึ่งพิงการนำเข้าทรัพยากรจากต่างประเทศเป็นหลัก ทำให้ต้องเสียค่าใช้จ่ายเป็นจำนวนมาก ดังนั้นความสามารถในการใช้และพัฒนาองค์ความรู้ทางด้านการเข้ารหัสสัญญาณวีดิทัศน์ และการสื่อสารผ่านเครือข่ายไร้สายจึงเป็นสิ่งสำคัญในการนำประเทศก้าวไปสู่การพึ่งพาตนเองทางด้านวิทยาศาสตร์และเทคโนโลยี

## 2. วัตถุประสงค์

- 2.1 เพื่อสร้างองค์ความรู้ทางกระบวนการเข้ารหัสสัญญาณวีดิทัศน์อัตราบิตต่ำ กรรมวิธีการเข้ารหัส สัญญาณวีดิทัศน์แยกส่วน การติดตามวัตถุโดยใช้เทคนิคทางคอมพิวเตอร์วิทัศน์ การจัดสรรบิตให้ กับวัตถุที่สนใจ รวมถึง เทคนิคการเพิ่มความทนทานของสัญญาณวีดิทัศน์เมื่อส่งผ่านช่องสัญญาณไร้สาย
- 2.2 เพื่อนำองค์ความรู้ที่ได้มาใช้ในการสร้างจริงงานประยุกต์เชิงวัตถุ เช่น ระบบการประชุมสัมมนาทางวีดิทัศน์

## 3. แนวเหตุผล และทฤษฎี

มาตรฐานการเข้ารหัสสัญญาณวีดิทัศน์ MPEG-4 [10.2] และ H.264/MPEG-4 Part 10 [10.3] ได้มีการกำหนดโครงสร้างการเข้าและถอดรหัส สัญญาณวีดิทัศน์ ซึ่งประกอบด้วย ส่วนการลงรหัสทรานส์ฟอร์ม (transform coding) การควอนไทซ์ (quantization) การลงรหัสเอนโทรปี (entropy coding) และการประมาณและชดเชยการเคลื่อนที่ (motion estimation and motion compensation) ในปัจจุบัน การเข้ารหัสเชิงวัตถุ และการส่งวีดิทัศน์ผ่านโครงข่ายไร้สายมีความสำคัญสำหรับงานประยุกต์ทางวีดิทัศน์ เช่น ระบบประชุมสัมมนาทางวีดิทัศน์ ระบบการประชุมเสมือน (virtual meeting) การส่งวีดิทัศน์ผ่านมือถือ (Video on Mobile) และ ระบบตรวจ

การณ์ทางไกล เป็นต้น      ในส่วนการเข้ารหัสเชิงวัตถุ กรรมวิธีการแยกส่วนภาพ (image segmentation) เป็นส่วนประกอบสำคัญ แต่อย่างไรก็ดี ไม่ได้มีการกำหนดกรรมวิธีไว้ในมาตรฐานการเข้ารหัสสัญญาณวิดีโอ ดังนั้น จึงมีการวิจัยทางนี้อย่างกว้างขวาง องค์ความรู้ที่สำคัญในการแยกส่วนภาพ คือ กรรมวิธีการประมวลผลภาพ ซึ่งประกอบด้วยทฤษฎีทางพื้นฐานทาง การตรวจจับขอบ การตรวจจับการเปลี่ยนแปลงระหว่างภาพ การหาค่าทางสถิติของลักษณะเด่นของภาพ เช่น สี พื้นผิว และ รูปร่าง และ กรรมวิธีทาง Morphology เป็นต้น ทั้งนี้ ในโครงการวิจัยที่ผ่านมา ได้นำเสนอกรรมวิธีการเข้ารหัสสัญญาณวิดีโอ แยกส่วน โดยใช้การแยกส่วนสัญญาณวิดีโอแบบทันทันโดยใช้วิธีการทางสถิติเชิงพารามตริกแบบปรับตัวได้ โดยปรับปรุงกรรมวิธีการลบฉากหลังให้มีประสิทธิภาพที่ดีขึ้น [10.4-10.7] และมีการนำเสนอกรรมวิธีการตรวจหาใบหน้าและมือมนุษย์ ซึ่งสามารถนำไปประยุกต์ใช้ได้กับงานประยุกต์การส่งวิดีโอทัศนภาษามือ [10.8-10.9] ในโครงการวิจัยปีนี้ จะเพิ่มประสิทธิภาพของกรรมวิธีการแยกส่วนภาพที่ได้พัฒนาขึ้น ซึ่งจะต้องมีความซับซ้อนต่ำ เพื่อให้เหมาะสมกับการนำไปประยุกต์ในส่วนตัวเข้ารหัสสัญญาณวิดีโอ เพื่อใช้ในการส่งสัญญาณวิดีโอเวลาจริง รวมถึง พัฒนาเทคนิคการตรวจหาและติดตามใบหน้า โดยใช้กล้องวิดีโอ 2 ตัว ในการส่งสัญญาณวิดีโอเวลาจริงผ่านโครงข่ายไร้สายนั้น ปัจจัยสำคัญ คือ ความน่าเชื่อถือของระบบการส่งสัญญาณวิดีโอเวลาจริง ซึ่งส่งผลถึง คุณภาพของสัญญาณวิดีโอที่ปลายทาง ว่ามีความผิดพลาดเป็นสิ่งที่สังเกตเห็นชัดเจนหรือไม่ ส่วนประกอบสำคัญที่เกี่ยวข้อง คือ กรรมวิธีการจัดสรรบิตของตัวเข้ารหัส และ เทคนิคในการเข้ารหัสเพื่อให้สัญญาณวิดีโอมีความทนทานต่อความผิดพลาดของ สัญญาณวิดีโอในการส่งผ่านโครงข่ายไร้สาย ในโครงการนี้จะได้ทำการศึกษาในส่วนนี้ เพื่อให้สัญญาณ วิดีโอทัศนมีคุณภาพที่ดี และมีความผิดพลาดน้อยที่สุด ซึ่งถือเป็นปัจจัยที่สำคัญต่อระบบการสื่อสารสัญญาณวิดีโอ

#### 4. ขั้นตอนการวิจัย

โครงการวิจัยช่วงปีที่ 5 นี้ได้ถูกแบ่งออกเป็นการทำงาน 2 ส่วน โดย จะเป็นการสานต่องานวิจัยทางด้านการเข้ารหัสสัญญาณวิดีโอที่อัตราบิตต่ำ จากโครงการในปีที่ 1 - 4 มีรายละเอียดของวิธีการทำงานในแต่ละส่วน และระยะเวลา ดังต่อไปนี้

ส่วนที่ 1 การสานต่อการพัฒนาเทคนิคการเข้ารหัสสัญญาณวิดีโอเชิงวัตถุที่มีประสิทธิภาพสูง มีความซับซ้อนน้อย และสามารถลดปริมาณข้อมูล ได้มากขึ้น ในส่วนของ การแยกส่วนภาพ และ การตรวจหา และ ติดตามใบหน้า

ส่วนที่ 2 การสานต่อส่วนการพัฒนาเทคนิคในการเข้ารหัสเพื่อให้สัญญาณวีดิทัศน์มีความทนทานต่อความ ผิดพลาดของสัญญาณวีดิทัศน์ ในการส่งผ่านโครงข่ายไร้สาย รวมถึง วิธีแก้ไข และปกปิดความผิดพลาด การเข้ารหัสแบบสเกล เพื่อเพิ่มคุณภาพของสัญญาณวีดิทัศน์ เพื่อทำการส่งบนโครงข่ายสื่อสารไร้สาย

**ทั้งนี้ มีแผนการดำเนินงานและระยะเวลาดังต่อไปนี้**

การดำเนินงาน	เดือน	เดือน	เดือน	เดือน
	ที่ 1-3	ที่ 4-6	ที่ 7-9	ที่ 10-12
1. ส่วนการพัฒนาเทคนิค การเข้ารหัสสัญญาณวีดิทัศน์เชิงวัตถุ	←————→			
2. ส่วนการพัฒนาเทคนิคในการเข้ารหัสเพื่อให้สัญญาณวีดิทัศน์มีความทนทานต่อความผิดพลาดของสัญญาณวีดิทัศน์ ในการส่งผ่านโครงข่ายไร้สาย	←————→			
3. ส่วนการทดสอบประสิทธิภาพของขั้นตอนวิธีที่นำเสนอ	←————→			
4. ส่วนการจัดทำรายงาน				←————→

## 5. ขอบเขต และเป้าหมาย

การวิจัยในโครงการจะครอบคลุมถึง

- 5.1 กระบวนการเข้ารหัสสัญญาณวีดิทัศน์อัตราบิตต่ำกรรมวิธีการเข้ารหัสสัญญาณวีดิทัศน์แยกส่วนการติดตามวัตถุโดยใช้เทคนิคทางคอมพิวเตอร์วีดิทัศน์การจัดสรรบิตให้กับวัตถุที่สนใจ รวมถึงเทคนิคการเพิ่มความทนทานของสัญญาณวีดิทัศน์เมื่อส่งผ่านช่องสัญญาณไร้สาย
- 5.2 การนำขั้นตอนวิธีของการเข้ารหัสที่พัฒนาได้ มาประยุกต์ควบคู่กับ วิธีแก้ไขและปกปิดความผิดพลาดเพื่อเพิ่มคุณภาพของสัญญาณวีดิทัศน์ เพื่อทำการส่งบนโครงข่ายสื่อสารไร้สาย ทั้งในแบบจำลองและสถานการณ์จริง
- 5.3 การทดสอบและวิเคราะห์ประสิทธิภาพขั้นตอนวิธีที่นำเสนอ

## 6. ส่วนของงานที่ได้ดำเนินการไปในปี 2549

- 6.1 นำเสนอกรรมวิธีการติดตามใบหน้าและการเคลื่อนที่ของบุคคลโดยใช้กล้องวีดิทัศน์ PTZ 2 ตัว พร้อมโปรแกรมต้นแบบ (รายละเอียดอยู่ในภาคผนวก ก)
- 6.2 ศึกษากรรมวิธีการติดตามใบหน้าและการเคลื่อนที่ของบุคคลโดยใช้กล้องวีดิทัศน์ PTZ มากกว่า 2 ตัว (รายละเอียดอยู่ในภาคผนวก ข)
- 6.3 นำเสนอกรรมวิธีการควบคุมอัตราใหม่สำหรับมาตรฐาน H.264/MPEG-4 Part 10 (รายละเอียดอยู่ในภาคผนวก ค)

## 7. ส่วนของงานที่จะดำเนินการต่อไป

- 7.1 นำเสนอวิธีการตรวจหา ติดตามและรู้จำใบหน้า โดยใช้กล้องวีดิทัศน์แบบสาย กัมเมย และ ซุ่ม 3 ตัวร่วมกันที่นำเสนอ เพื่อให้สามารถประยุกต์ใช้ได้สถานการณ์จริง
- 7.2 พัฒนาโมเดลการควบคุมอัตราใหม่สำหรับมาตรฐาน H.264/MPEG-4 Part 10 เพื่อการส่งผ่านช่องสัญญาณไร้สายที่มีเวลาประวิงต่ำ

## 8. ประโยชน์ของงานวิจัย

- 8.1 พัฒนาองค์ความรู้ด้านการเข้ารหัสสัญญาณวีดิทัศน์เชิงวัตถุและการส่งผ่านโครงข่ายไร้สาย
- 8.2 พัฒนาระบบที่ระดับปริญญาโทและเอกที่มีทักษะ ความรู้ความสามารถ และมีความรับผิดชอบในหน้าที่
- 8.3 พัฒนาศักยภาพของภาควิชาให้ไปสู่การเป็นผู้นำทางศาสตร์การสื่อสารมัลติมีเดีย

## 9. ผลผลิตที่ได้จากงานวิจัยปี 2549 เพื่อใช้เป็นดัชนีชี้วัดความสัมฤทธิ์ผลของโครงการ

- 9.1 องค์ความรู้ด้านการเข้ารหัสสัญญาณวีดิทัศน์ การตรวจหาและติดตามใบหน้า และด้านการเชื่อมต่อโครงข่ายไร้สาย
- 9.2 บทความทางวิชาการระดับนานาชาติ และระดับชาติ จำนวน 9 บทความ ได้แก่
  - 9.2.1 พิชัย อำนวยกาญจนสิน สุภาวดี อร่ามวิทย์ และ ธนารัตน์ ชลิดาพงศ์ “การระบุใบหน้าบุคคลโดยใช้กล้อง PTZ สองตัวทำงานร่วมกัน” การประชุมทางวิศวกรรมไฟฟ้าครั้งที่ 29 พัทยา ชลบุรี พฤศจิกายน 2549
  - 9.2.2 นงลักษณ์ เขียมจรัส และ สุภาวดี อร่ามวิทย์ “การปรับปรุงระเบียบวิธีควบคุมอัตราสำหรับมาตรฐานการเข้ารหัสสัญญาณวีดิทัศน์ H.264 โดยใช้แบบจำลองอัตราบิดและความเพี้ยนแบบโคชี” การประชุมทางวิศวกรรมไฟฟ้าครั้งที่ 29 พัทยา ชลบุรี พฤศจิกายน 2549
  - 9.2.3 วิสูตร หาญทองค์ และ สุภาวดี อร่ามวิทย์ “การจัดเรียงแมโครบล็อกแบบยืดหยุ่นร่วมกับการเข้ารหัสแก้ไขความผิดพลาดไปหน้าสำหรับการส่งวีดิทัศน์ไร้สาย H.264” การประชุมทางวิศวกรรมไฟฟ้าครั้งที่ 29 พัทยา ชลบุรี พฤศจิกายน 2549
  - 9.2.4 N. Eiamjumrus and S. Aramvith, “Cauchy Based Rate-Distortion Optimization Model for H.264 Rate Control,” *Proceeding of IEEE Asia*

*Pacific Conference on Circuits and Systems (APCCAS) 2006*, Singapore, December 2006.

9.2.5 P. Amnuaykanchanasin, S. Aramvith, and T.H. Chalidabhongse, "Real- Time Face Identification using Two Cooperative Active Cameras," *Proceeding of IEEE International Conference on Control, Automation, Robotics and Vision*

*(ICARCV) 2006*, Singapore, December 2006.

9.2.6 N. Eiamjumrus and S. Aramvith, "New Rate-Control Scheme based on Cauchy Rate-Distortion Optimization Model for H.264 Video Coding," *Proceeding of IEEE International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS) 2006*, Japan, December 2006.

9.2.7 S. Aramvith and W. Hantanong, "Joint Flexible Macroblock Ordering and FEC for H.264 Wireless Video Transmission," *Proceeding of IEEE International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS) 2006*, Japan, December 2006.

9.2.8 N. Eiamjumrus and S. Aramvith, "Rate-Control Scheme based on Cauchy Rate-Distortion Optimization Model for H.264 under Low Delay Constraint," *Proceeding of IEEE International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP) 2006*, USA, December 2006.

9.3 โปรแกรมต้นแบบการตรวจหาและติดตามใบหน้า โดยใช้กล้อง

9.4 ซอฟต์แวร์จำลองการรับ-ส่ง H.264/MPEG-4 part 10 ที่ใช้ร่วมกับตัวเข้ารหัส-ถอดรหัสวีดิทัศน์มาตรฐาน H.264 AVC/MPEG-4 Part 10

## 10. เอกสารอ้างอิง

- 10.1 ITU-T Draft recommendation H.263, "Video Coding for low bit-rate communication," May 1997.
- 10.2 ISO/IEC 14496, "Information Technology – Coding of audio-visual objects".
- 10.3 T. Wiegand, G. Sullivan and A. Luthra, "Draft ITU-T commendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC)", Joint Video Team, Doc. JVT-G050, Pattaya (Thailand), March 2003.
- 10.4 T. Thongkamwitoon, S. Aramvith, and T. H. Chalidabhongse, "An Adaptive Real-Time Background Substraction and Moving Shadows Detection," *Proceeding of IEEE International Conference on Multimedia and Expo (ICME 2004)*, Taipei, Taiwan, June 2004.
- 10.5 ธิรพิรุฬห์ ทองคำวิฑูรย์ สุภาวดี อร่ามวิทย์ และ ธนารัตน์ ชลิดาพงศ์ "ระเบียบวิธีการลบฉากหลังเวลาจริงแบบปรับตัวได้สำหรับการแยกส่วนภาพ วีดิทัศน์" การประชุมทาง วิศวกรรมไฟฟ้าครั้งที่ 27 ขอนแก่น พฤศจิกายน 2547
- 10.6 T. Thongkamwitoon, S. Aramvith, and T. H. Chalidabhongse, "Adaptive Background Subtraction Algorithm with Auto Brightness Control for Consumer-type Cameras," *Proceeding of International Workshop on Advanced Image Technology (IWAIT) 2005*, Jeju, Korea, January 2005.
- 10.7 T. Thongkamwitoon, S. Aramvith, and T. H. Chalidabhongse, "Non-Linear Learning Control Factor for Statistical Adaptive Background Subtraction Algorithm," *Proceeding of*



International Symposium on Circuits and Systems (ISCAS)  
2005, Kobe, Japan, May 2005.

10.8 N. Soontranon, S. Aramvith, and T. H. Chalidabhongse, "Face and Hands Localization and Tracking for Sign Language Recognition," *Proceeding of International Symposium on Communication and Information Technologies (ISCIT2004)* , Sapporo, Japan, October 2004.

10.9 N. Soontranon, S. Aramvith, and T. H. Chalidabhongse, "Improved Face and Hand Tracking for Sign Language Recognition," To appear, *Proceeding of IEEE International Conference on Information Technology (ITCC 2005)*, Las Vegas, April 2005.

#### 11. รายชื่อผู้ทรงคุณวุฒิ พร้อมสถานที่การทำงาน

11.1 รศ. ดร. มนต์ สังวรศิลป์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

11.2 ผศ. ดร. วุฒิพงษ์ อารีกุล มหาวิทยาลัยเกษตรศาสตร์

สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย

## ภาคผนวก ก

### 1. จุดประสงค์ของงานวิจัย

- 1.1 พัฒนาระบบและโปรแกรมประยุกต์การติดตามใบหน้าด้วยกล้องวีดิทัศน์ PTZ 2 ตัว ให้ทำงานร่วมกัน เพื่อประยุกต์ใช้ในงานด้านการสอดส่องรักษาความปลอดภัย
- 1.2 พัฒนาส่วนการระบุบุคคล เพื่อเพิ่มประสิทธิภาพในการระบุบุคคลด้วยการจับคู่ใบหน้าให้มีความถูกต้องในการระบุบุคคลมากขึ้น
- 1.3 หาแนวทางและขั้นตอนวิธีใหม่สำหรับมาประยุกต์ใช้ในการระบุบุคคลร่วมกับขั้นตอนวิธีการจับคู่ใบหน้า
- 1.4 พัฒนาขั้นตอนวิธีการติดตามบุคคลในส่วนของงานการตรวจหาและติดตามบุคคลต่อเนื่องระหว่างกล้องวีดิทัศน์ ให้มีประสิทธิภาพในการติดตามบุคคลเพิ่มขึ้น และสามารถติดตามบุคคลพร้อมกันมากกว่าหนึ่งคน
- 1.5 สามารถนำประโยชน์ที่ได้จากการใช้กล้องวีดิทัศน์ 2 ตัว ให้ทำงานร่วมกัน มาเพิ่มประสิทธิภาพในกระบวนการ รู้จำใบหน้าและระบุบุคคลระหว่างกล้อง
- 1.6 เพื่อเป็นแนวทางในการนำระบบไปประยุกต์ใช้กับลักษณะงานเป้าหมายในสถานการณ์จริงได้

### 2. รายละเอียดโดยย่อของงานวิจัย

งานวิจัยนี้นำเสนอระบบกล้องหลายตัวที่สามารถนำไปประยุกต์ใช้งานสำหรับระบบสอดส่องรักษาความปลอดภัยภายในอาคาร โดยระบบจะประกอบด้วยกล้องวีดิทัศน์ PTZ 2 ตัวซึ่งจะมีการประมวลผลร่วมกันเพื่อประโยชน์ในส่วนการตรวจหาติดตาม และส่วนของการระบุบุคคล โดยใช้วิธีการจับคู่ใบหน้า งานวิจัยนี้จะใช้ประโยชน์ของระบบกล้องหลายตัวและความสามารถของกล้องวีดิทัศน์ PTZ กล่าวคือ ความสามารถในการส่ายและก้มเงยเปลี่ยนมุมมองการรับภาพและการซูมของกล้องวีดิทัศน์ PTZ เพื่อให้ได้ตำแหน่งใบหน้าที่ดีที่สุด ประโยชน์ของระบบกล้องหลายตัว ที่จะมีมุมมองการรับภาพที่กว้างกว่า ทำให้เพิ่มโอกาสในการประมวลผลกับภาพที่มีความเหมาะสมมากที่สุด

ซึ่งงานวิจัยในภาคนี้จะมุ่งเน้นไปที่การพัฒนาขั้นตอนวิธีในแต่ละองค์ประกอบของ แต่ละส่วนการทำงานของระบบ และการวิเคราะห์ประสิทธิภาพของระบบในลักษณะการ เปรียบเทียบผลกระทบที่เกิดขึ้นเมื่อมีการปรับเปลี่ยนในแต่ละขั้นตอนวิธี ส่วนการทำงานที่มีความสำคัญและเป็นส่วนหลักในการพัฒนาปรับเปลี่ยน คือ ส่วนของการระบุบุคคล ซึ่งในภาคที่ ผ่านมาได้เลือกใช้การจับคู่ภาพใบหน้าโดยใช้ระยะ Bhattacharyya ในการเปรียบเทียบภาพรับเข้า และภาพใบหน้าในฐานข้อมูล โดยข้อมูลที่ใช้ คือ ข้อมูลการกระจายตัวของสี หรือ ฮิสโตแกรมของ ภาพใบหน้า ซึ่งข้อดีของลักษณะนี้ คือ มีความซับซ้อนต่ำและประสิทธิภาพในด้านเวลาในการ ประมวลผล แต่ข้อเสีย คือ ประสิทธิภาพในการระบุบุคคลไม่มาก เนื่องจากไม่มีการวิเคราะห์เชิง พื้นที่ จึงเป็นแนวความคิดของการพัฒนาในส่วนนี้ โดยจะเพิ่มการใช้ข้อมูลเชิงพื้นที่ของภาพใบหน้า สำหรับการจับคู่ใบหน้า ซึ่งจะแบ่งได้เป็นสองลักษณะ คือ

- แบ่งพื้นที่ในการจับคู่ภาพ กล่าวคือ จะมีการแบ่งบริเวณของภาพใบหน้าที่ได้ ออกเป็นส่วน ๆ แล้วทำการหาฮิสโตแกรมสีในแต่ละส่วน จากนั้นจะเปรียบเทียบตามตำแหน่ง
- ตรวจสอบตำแหน่งดวงตา ปากและจมูก เพื่อใช้เป็นลักษณะในการจับคู่ ซึ่งน่าจะ ส่งผลต่อประสิทธิภาพในการระบุใบหน้า

สำหรับรูปแบบการทำงานอีกส่วนของระบบ คือ การตรวจหาและติดตามบุคคล อย่างต่อเนื่องระหว่างกล้องวีดิทัศน์ 2 ตัว ซึ่งจะใช้ประโยชน์ของระบบกล้องหลายตัว ในลักษณะ ของการส่งผ่านข้อมูลบุคคลระหว่างกล้องวีดิทัศน์ตัวหนึ่งไปยังกล้องวีดิทัศน์อีกตัวหนึ่ง โดยใน งานวิจัยนี้จะเสนอแนวทางการประมวลผลเพื่อพัฒนาในส่วนของการติดตามบุคคลอย่างมี ประสิทธิภาพ โดยสามารถรองรับการติดตามบุคคลภายในบริเวณภาพได้มากกว่า 1 คน

### 3. เนื้อหาและผลการทดลอง

จากงานวิจัยในปีการศึกษาที่ผ่านมา ได้สร้างระบบสอดส่องรักษาความปลอดภัย ที่มีการใช้กล้องวีดิทัศน์ 2 ตัวให้มีการประมวลผลและทำงานร่วมกัน โดยแบ่งออกได้เป็น 2 รูปแบบ หลัก คือ

1. การประมวลผลร่วมกันสำหรับระบุบุคคล
2. การทำงานร่วมกันสำหรับการตรวจหาติดตามใบหน้าอย่างต่อเนื่องระหว่างกล้องวีดิทัศน์ 2 ตัว

ซึ่งจากผลการวิจัยที่ผ่านมา สิ่งหนึ่งที่ได้รับและเกิดขึ้น ก็คือ ระบบต้นแบบของรูปแบบงาน แต่ละส่วน แต่จากผลการทดลองที่ผ่านมาจะเห็นว่า ยังมีขั้นตอนวิธีในหลาย ๆ ส่วน ที่มีปัญหา ยัง

ผลโดยตรงต่อคุณภาพและประสิทธิภาพของระบบ ดังนั้นงานวิจัยในปีการศึกษานี้จึงมุ่งเน้นไปที่การพัฒนาเพื่อปรับปรุง แก้ไข เพื่อเพิ่มประสิทธิภาพและความถูกต้องของระบบ รวมทั้งนำขั้นตอนวิธีอื่นมาประยุกต์ใช้ โดยคำนึงถึงความเหมาะสม และความเป็นไปได้ที่จะประยุกต์ใช้ให้เข้ากับรูปแบบของระบบก่อนหน้า

ปัญหา แนวทางการปรับปรุงและพัฒนา และผลการทดลอง ของรูปแบบงานในแต่ละส่วน มีดังนี้

### 3.1. ส่วนการประมวลผลร่วมกันสำหรับระบุบุคคล

#### 3.1.1 ปัญหา

จุดมุ่งหมายของระบบต้นแบบของการประมวลผลร่วมกันสำหรับระบุบุคคล คือ การที่ระบบสามารถระบุบุคคลได้ว่าบุคคล ณ ขณะนั้นที่ระบบตรวจหาได้ คือ ใครในฐานข้อมูลบุคคลของระบบ โดยในการทดลองจะไม่มีข้อกำหนดวิธีการเดินของบุคคลเป้าหมายที่แน่นอน ทำให้ตำแหน่งทิศทางของใบหน้าของบุคคลที่ระบบตรวจหาได้ไม่มีความแน่นอน ซึ่งส่งผลกระทบต่อขั้นตอนการระบุบุคคล เพราะโดยทั่วไปการระบุบุคคล ทิศทางของใบหน้าส่งผลอย่างมากต่อความถูกต้องของการระบุบุคคล ซึ่งจากงานวิจัยของรูปแบบงานในลักษณะนี้ที่ผ่านมา ใบหน้าตรงเป็นตำแหน่งทิศทางที่ให้ความถูกต้องของการระบุบุคคลมากที่สุด

จะเห็นว่างานวิจัยในปีการศึกษาที่ผ่านมา ได้ใช้การจับคู่บุคคล โดยใช้ขั้นตอนวิธีการวัดความคล้าย (Similarity measure) ด้วยการกระจายตัวของสีของส่วนบริเวณใบหน้า เป็นขั้นตอนวิธีในการระบุบุคคล เนื่องด้วยเหตุผลของความไม่แน่นอนของทิศทางใบหน้าที่ตรวจหาได้ (กำหนดทิศทางของใบหน้าบุคคลที่ตรวจหาได้ เป็น 3 ทิศทาง ใบหน้าตรง ใบหน้าด้านเฉียงซ้าย ใบหน้าด้านเฉียงขวา) และต้องคำนึงถึงเวลาที่ใช้ในการประมวลผลต้องรองรับการทำงานในเวลาจริงด้วย ซึ่งเมื่อพิจารณาที่ผลการทดลองประสิทธิภาพความถูกต้องของการระบุใบหน้าจะอยู่ที่ประมาณ 60-70 เปอร์เซ็นต์ ดังนั้นแนวทางในงานวิจัยสำหรับส่วนนี้ คือ ปรับปรุงขั้นตอนวิธีในการระบุบุคคล เพื่อเพิ่มประสิทธิภาพในการระบุบุคคลให้มีความถูกต้องมากขึ้น และเป็นแนวทางในการนำไปพัฒนาต่อไป

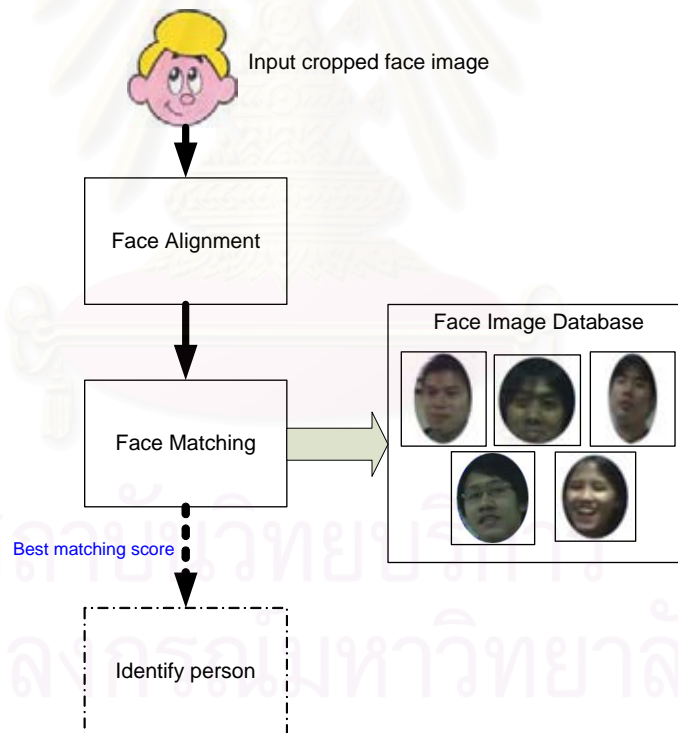
#### 3.1.2 แนวทางการปรับปรุงและพัฒนา

สำหรับงานวิจัยในปีการศึกษานี้จะแบ่งการวิจัยเพื่อปรับปรุงและพัฒนาเป็น 2 รูปแบบ คือ

1. ศึกษาและทดลอง รูปแบบของการจับคู่ใบหน้า ในเชิงพื้นที่
2. ศึกษาและทดลอง การตรวจหาลักษณะเด่นบนใบหน้า เพื่อเป็นแนวทางในการนำไปประยุกต์ใช้ในการระบุบุคคลในงานวิจัยขั้นถัดไป

### 3.1.2.1 รูปแบบของการจับคู่ใบหน้าในเชิงพื้นที่

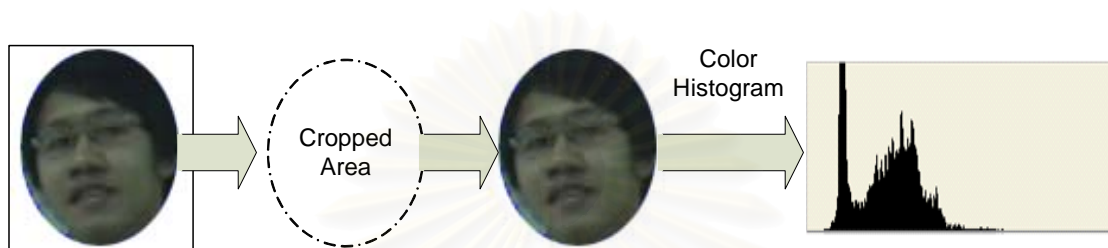
กระบวนการระบุใบหน้าของงานวิจัยในปีการศึกษาที่ผ่านมาแสดงได้ดังรูปที่ 1 ในที่นี้จะมุ่งพิจารณาปรับปรุงพัฒนาในส่วนของขั้นตอนวิธีการจับคู่ภาพใบหน้า (Face Matching) ซึ่งวิธีการเดิมที่ใช้ คือ จะใช้หลักการการวัดความคล้าย (Similarity measure) ระหว่างภาพใบหน้าขาเข้า และภาพใบหน้าที่เกิดขึ้นไว้ในฐานข้อมูล (Face image database) โดยใช้สัมประสิทธิ์ Bhattacharyya เป็นตัวบ่งชี้ค่าความคล้ายของภาพใบหน้าแต่ละคู่ภาพ โดยลักษณะ (Feature) ที่นำมาใช้วัด คือ ค่าการกระจายตัวของสีหรือฮิสโตแกรมสี (Color histogram) ของบริเวณใบหน้าที่ผ่านการตัดส่วนภาพเป็นรูปวงรีดังตัวอย่างในรูปที่ 1



รูปที่ 1 โครงสร้างกระบวนการระบุใบหน้า

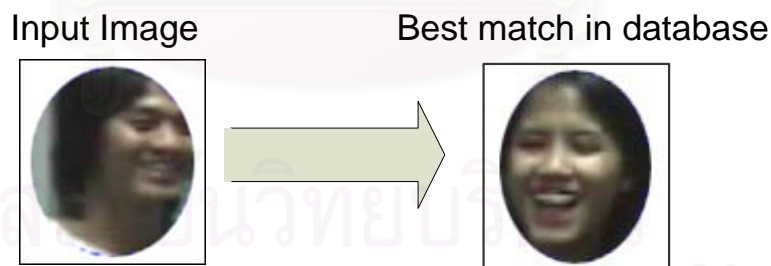
โดยค่าฮิสโตแกรมสีที่ถูกนำมาใช้ คือ ค่าฮิสโตแกรมสีโดยรวมของบริเวณใบหน้าที่ถูกตัดส่วนภาพมาได้ แสดงได้ดังรูปที่ 2 ซึ่งเมื่อพิจารณาแล้วจะเห็นว่าปัญหาของการจับคู่จะเกิดขึ้นในกรณีที่พิจารณาภาพในลักษณะโดยรวม (Global) และลักษณะท้องถิ่น (Local) จากตัวอย่าง รูปที่

4 จะเห็นว่าทุกภาพวงรีมีรูปแบบหรือลักษณะพื้นผิวที่แตกต่างกัน แต่ถ้านำมาพิจารณาในเชิงค่าฮิสโตแกรมสีแล้ว ทุกภาพวงรีจะมีค่าเท่ากันหมด คือ มีค่าสีขาว และค่าสีดำ อย่างละครึ่งภาพ เพราะจำนวนบริเวณหรือจุดการเกิดสี (ดำหรือขาว) มีค่าเท่ากันหมด แม้ตำแหน่งการเกิดสีนั้น ๆ จะต่างกัน แต่ก็ไม่มีผลต่อการคิดหาค่าฮิสโตแกรมในลักษณะโดยรวม

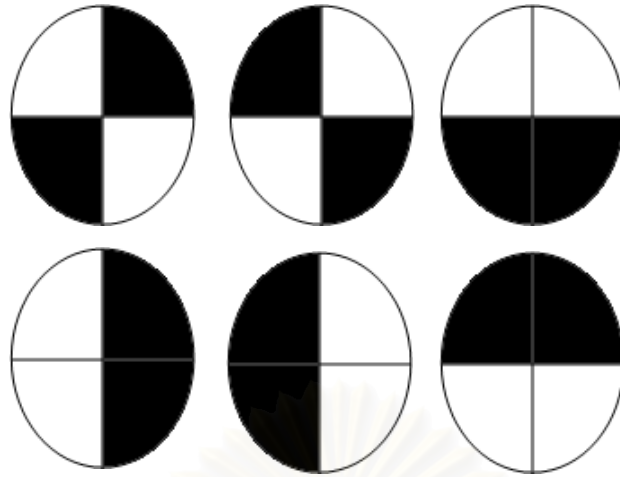


รูปที่ 2 ขั้นตอนวิธีการหาฮิสโตแกรมสีแบบเดิม

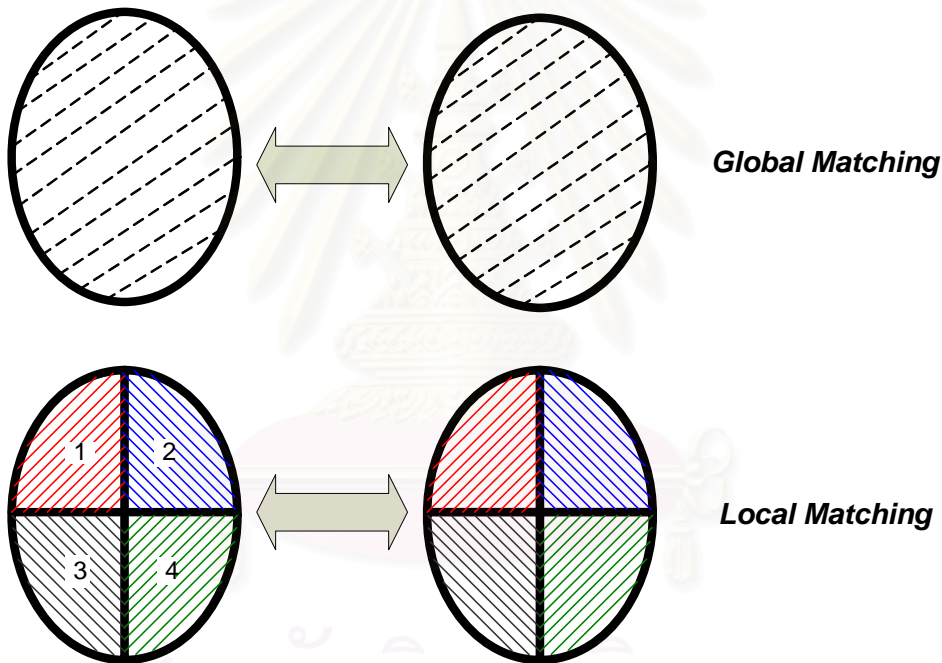
รูปที่ 3 แสดงตัวอย่างของการจับคู่ที่ผิดพลาด จากรูปจะสังเกตเห็นได้ว่า ถ้าพิจารณาเฉพาะส่วนบริเวณสีดำในรูปภาพใบหน้าทั้งสองจะค่อนข้างมีพื้นที่ใกล้เคียงกัน ซึ่งส่งผลให้ระบบจำแนกจับคู่ได้ใกล้เคียงกัน แต่ในความเป็นจริงทั้งสองรูปมีตำแหน่งใบหน้าที่แตกต่างกันอย่างเห็นได้ชัด



รูปที่ 3 การจับคู่ภาพใบหน้าผิดพลาด



รูปที่ 4 ตัวอย่างการพิจารณาฮิสโตแกรมสีของภาพที่เกิดความผิดพลาด



รูปที่ 5 แนวทางการแก้ปัญหาในการหาฮิสโตแกรมสีเชิงพื้นที่

ดังนั้นแนวทางการแก้ปัญหาในงานวิจัย คือ การหาค่าฮิสโตแกรมสีที่พิจารณาเชิงพื้นที่ด้วย กล่าวคือ การแบ่งพื้นที่การหาค่าฮิสโตแกรมสี กล่าวคือ จะเปลี่ยนแนวทางการพิจารณาในลักษณะโดยรวม (Global) มาเป็นการพิจารณาในลักษณะท้องถิ่น(Local) โดยในงานวิจัยนี้จะแบ่งพื้นที่การพิจารณาภาพใบหน้าเป็น 4 ส่วน จากนั้นจะหาค่าฮิสโตแกรมสีแยกในแต่ละส่วน รวมทั้งการจับคู่ก็จะแยกในแต่ละส่วนบริเวณ เพราะฉะนั้นสามารถคำนวณหาคะแนนการจับคู่ใบหน้าด้วยระยะ Bhattacharyya (Bhattacharyya Distance) ของใบหน้าขาเข้า เมื่อเทียบกับภาพใบหน้าแต่ละภาพในฐานข้อมูล ได้ดังสมการที่ 1

$$Bh_T = \frac{Bh_{TL} + Bh_{TR} + Bh_{BL} + Bh_{BR}}{4} \quad (1)$$

เมื่อ  $Bh_T$  คือ ระยะ Bhattacharyya รวมของคู่ภาพนั้น ๆ

$Bh_{TL}$  คือ ระยะ Bhattacharyya เปรียบเทียบระหว่างฮิสโตแกรมบริเวณที่ 1 ในรูปที่

5

$Bh_{TR}$  คือ ระยะ Bhattacharyya เปรียบเทียบระหว่างฮิสโตแกรมบริเวณที่ 2 ในรูปที่

5

$Bh_{BL}$  คือ ระยะ Bhattacharyya เปรียบเทียบระหว่างฮิสโตแกรมบริเวณที่ 3 ในรูปที่

5

$Bh_{BR}$  คือ ระยะ Bhattacharyya เปรียบเทียบระหว่างฮิสโตแกรมบริเวณที่ 4 ในรูปที่

5

โดยที่ ระยะ Bhattacharyya สามารถคำนวณหาได้ ดังสมการที่ 2

$$Bh = \sqrt{1 - \rho[\hat{p}, \hat{q}]} \quad (2)$$

เมื่อ  $\rho[\hat{p}, \hat{q}] = \sum_{u=1}^m \sqrt{\hat{p}_u \times \hat{q}_u}$  เป็นค่าสัมประสิทธิ์ Bhattacharyya  $\hat{p}$  เป็นฮิสโตแกรมสี

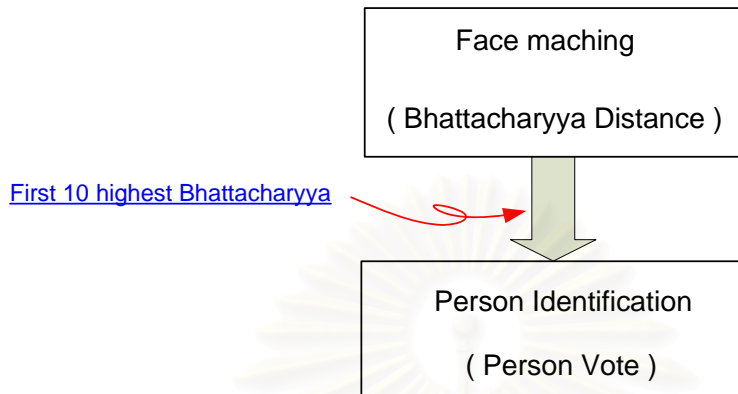
ของภาพใบหน้าในฐานข้อมูล และ  $\hat{q}$  เป็นฮิสโตแกรมสีของภาพใบหน้าขาเข้าที่ต้องการนำมา  
เปรียบเทียบ โดย  $\hat{p} = \{\hat{p}_u\}_{u=1, \dots, m}$  (เมื่อ  $\sum_{u=1}^m \hat{p}_u = 1$ ) และ  $\hat{q} = \{\hat{q}_u\}_{u=1, \dots, m}$  (เมื่อ  $\sum_{u=1}^m \hat{q}_u = 1$ ) เมื่อ  $m$   
เป็นจำนวนถังสีของฮิสโตแกรม (bin histogram)

### 3.1.2.1.1 ผลการทดลองและวิเคราะห์

ในการทดลองวัดประสิทธิภาพ จะทำการทดลองกับฐานข้อมูลใบหน้าตรงทั้งหมด 200  
ภาพ โดยวิธีการทดลอง คือ นำภาพใบหน้าแต่ละภาพในฐานข้อมูลมาทำการจับคู่ภาพใบหน้าใน  
ฐานข้อมูลทั้งหมด ผลของการจับคู่ใบหน้าไม่จำเป็นต้องได้ภาพใบหน้านั้นเสมอ เพราะในขั้นตอน  
การระบุบุคคลส่วนสุดท้าย จะมีส่วนของการโหวตบุคคล ดังกระบวนการรูปที่ 6 เหตุผลคือ  
เนื่องจากภาพใบหน้าที่ตรวจหาได้ แม้จะสามารถจำแนกได้คร่าว ๆ ว่าเป็นใบหน้าตรง หรือใบหน้า



เอียงซ้ายหรือขวา แต่ก็ยังไม่มีคะแนนในตำแหน่งและทิศทางของรูปทรงใบหน้าที่ได้ การโหวตจาก 10 อันดับที่มีค่าคะแนนการจับคู่สูงที่สุด จึงถูกนำมาประยุกต์ใช้



รูปที่ 6 ขั้นตอนในการระบุบุคคล

ในการทดลองสิ่งหนึ่งที่ถูกนำมาพิจารณาด้วย คือ ปริภูมิสีที่เลือกใช้ในการนำมาหาฮิสโตแกรมสี โดยในการทดลองจะเลือกใช้ 2 แบบ คือ ปริภูมิสี RGB และ ภาพระดับเทา โดยจะเป็นตัวชี้ความน่าจะเป็นในการเลือกใช้อองค์ประกอบสีที่นำมาหาฮิสโตแกรม

จากผลการทดลองทั้งในตารางที่ 1 และ 2 จะเห็นว่า การระบุบุคคลโดยใช้การจับคู่ฮิสโตแกรมสีที่แยกตามพื้นที่ จะให้ค่าเปอร์เซ็นต์ความผิดพลาดที่ลดลงทุกองค์ประกอบที่ใช้หาฮิสโตแกรม ซึ่งแสดงให้เห็นว่าการแบ่งพื้นที่จะช่วยลดเปอร์เซ็นต์ความผิดพลาด ซึ่งก็ส่งผลให้เปอร์เซ็นต์ความถูกต้องเพิ่มขึ้นด้วย ในส่วนของการเลือกใช้อปริภูมิสี แม้ว่าจากผลการทดลองในตารางที่ 1 ผลของการใช้ฮิสโตแกรมขององค์ประกอบสีเดียว จะให้ค่าความผิดพลาดน้อยที่สุด แต่จะเห็นว่าการใช้อองค์ประกอบสีแดงและน้ำเงินให้ค่าเปอร์เซ็นต์ความผิดพลาดที่มากกว่า และยังมากกว่าผลของการระบุบุคคลโดยใช้ฮิสโตแกรมของภาพระดับเทา เนื่องจากว่าองค์ประกอบสีทั้งสามของปริภูมิสี RGB ไม่มีความสัมพันธ์ระหว่างกัน ดังนั้นการที่นำผลการระบุบุคคลทั้งสามมาคิดคำนวณร่วมกัน จึงไม่มีความสัมพันธ์ที่จะแสดงได้ ต่างกับการใช้ฮิสโตแกรมของภาพระดับเทา ซึ่งคำนวณมาจากค่าความเข้มขององค์ประกอบในปริภูมิสี RGB ดังนั้นถ้าพิจารณาเฉพาะผลในการทดลองนี้ การใช้ภาพระดับเทา ควบคู่ไปกับการจับคู่ใบหน้าแยกตามพื้นที่ น่าจะเหมาะสมและช่วยเพิ่มเปอร์เซ็นต์ความถูกต้องของการระบุบุคคลให้มากขึ้น

ตารางที่ 1 เปรียบเทียบเปอร์เซ็นต์ความผิดพลาดในการระบุบุคคลเมื่อใช้ปริภูมิสี RGB

องค์ประกอบสี (ที่ใช้ในการหาฮิสโตแกรมสี)	เปอร์เซ็นต์ความผิดพลาดในการระบุบุคคล แยกตามขั้นตอนวิธีการจับคู่ภาพใบหน้า (%)	
	ใช้ฮิสโตแกรมรวม	ใช้ฮิสโตแกรมแยกตามพื้นที่
องค์ประกอบสีแดง	30	22.5
องค์ประกอบสีเขียว	21.5	12.5
องค์ประกอบสีน้ำเงิน	31.5	20.5

ตารางที่ 2 เปรียบเทียบเปอร์เซ็นต์ความผิดพลาดในการระบุบุคคลเมื่อใช้ภาพระดับเทา

องค์ประกอบสี (ที่ใช้ในการหาฮิสโตแกรมสี)	เปอร์เซ็นต์ความผิดพลาดในการระบุบุคคล แยกตามขั้นตอนวิธีการจับคู่ภาพใบหน้า (%)	
	ใช้ฮิสโตแกรมรวม	ใช้ฮิสโตแกรมแยกตามพื้นที่
ภาพระดับเทา	26.5	16.5

### 3.1.2.2 การตรวจหาลักษณะเด่นบนใบหน้า (Facial feature)

ในหัวข้อที่ผ่านมาเป็นกระบวนการในการพัฒนาและเพิ่มประสิทธิภาพในการระบุใบหน้า โดยใช้วิธีการจับคู่ใบหน้าด้วยฮิสโตแกรมสี ซึ่งแม้ว่าจะช่วยเพิ่มประสิทธิภาพความถูกต้องในการระบุบุคคลในระดับหนึ่ง แต่ระบบก็จำเป็นต้องหาลักษณะเพิ่มเติมเพื่อมาช่วยเสริมให้การระบุบุคคลมีประสิทธิภาพมากยิ่งขึ้น โดยลักษณะที่เพิ่มเข้ามายังต้องมีความเหมาะสมที่จะสามารถประยุกต์เข้ากับระบบที่สร้างได้ ทั้งในด้านเวลาในการประมวลผลและรูปแบบหรือข้อจำกัดของระบบ

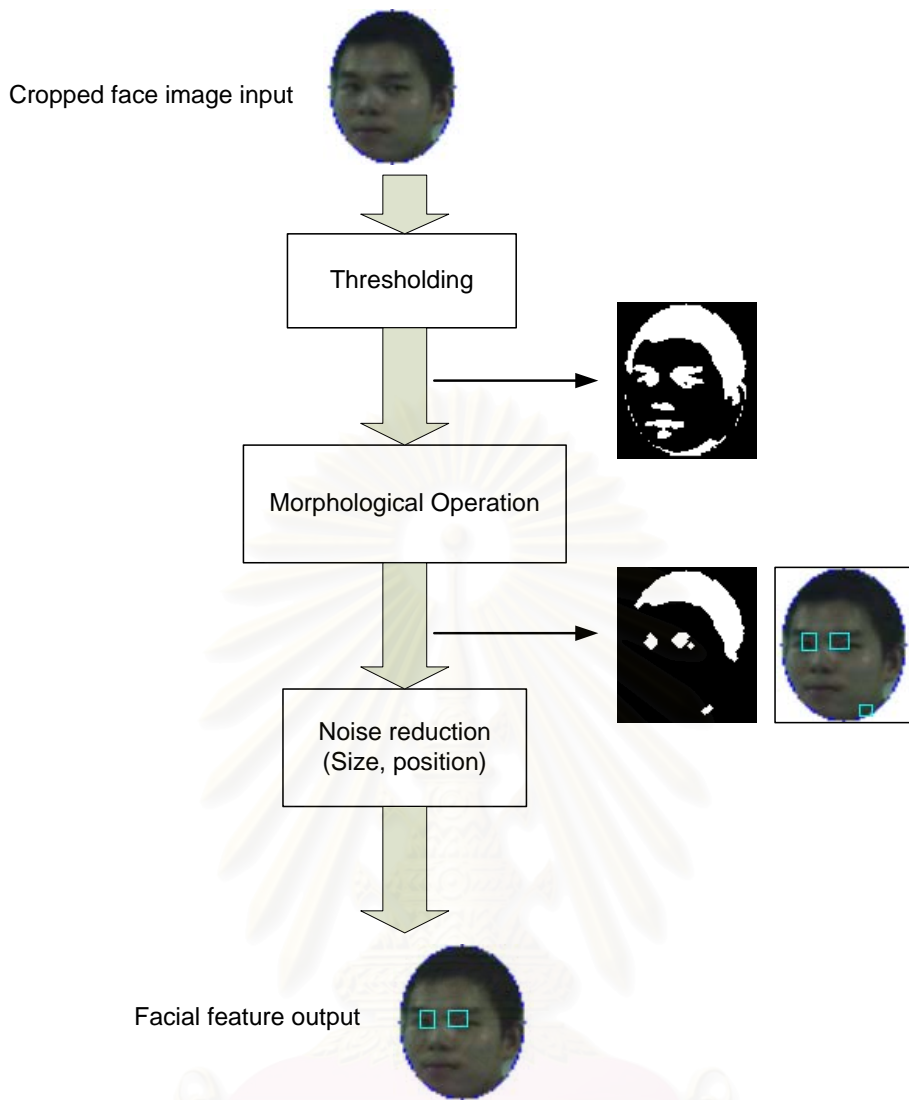
รายละเอียดของใบหน้า จึงเป็นลักษณะหนึ่งที่มีความเป็นไปได้ และได้ถูกนำมาทดลองเบื้องต้นในงานวิจัยและรายงานนี้ เพื่อเป็นแนวทางในการนำไปประยุกต์ใช้ในการระบุบุคคลในงานวิจัยขั้นถัดไป โดยทั่วไป รายละเอียดของใบหน้าที่สำคัญ คือ ตำแหน่งดวงตา จมูก และปาก

รวมถึงระยะห่างของอวัยวะแต่ละส่วนบนใบหน้า ซึ่งในงานวิจัยส่วนใหญ่ก็จะเป็นการตรวจหาในมุมมองใบหน้าด้านตรง ซึ่งจะสามารถวิเคราะห์หาลักษณะที่ต่างกันของใบหน้าได้ชัดเจน แต่จะเห็นว่าจากผลการตรวจหาที่ได้ของระบบ ทำให้ขั้นตอนวิธีบางอย่างไม่เหมาะสมกับการประยุกต์ใช้ ดังนั้นในงานวิจัยจึงได้ทดลองทดสอบการหาองค์ประกอบรายละเอียดของใบหน้า เพื่อนำมาพิจารณาความเป็นไปได้ ข้อผิดพลาด รวมทั้งแนวทางในการปรับปรุงในส่วนงานถัดไป

กระบวนการทำงานแสดงได้ดังรูปที่ 7 โดยเริ่มจากภาพขาเข้าจะถูกแปลงเป็นภาพฐานสอง (Binary image) โดยกำหนดเป็นค่าขีดเริ่มเปลี่ยนที่ใช้ในการจำแนกระหว่างบริเวณที่เป็นสีผิว (สีอ่อน) และบริเวณที่เป็นรายละเอียดใบหน้า (สีเข้ม) จากนั้นจะผ่านกระบวนการ morphological เพื่อกำจัดกลุ่มบริเวณที่ไม่ต้องการ ซึ่งส่วนใหญ่ จะเป็นกลุ่มบริเวณที่มีขนาดเล็ก ๆ ซึ่งหลังจากผ่านกระบวนการนี้ จะเหลือแต่กลุ่มบริเวณสำคัญภายในบริเวณใบหน้า ซึ่งจะใช้คุณลักษณะ 2 แบบ ในการระบุตำแหน่งรายละเอียดสำคัญบนใบหน้า คือ

- **ขนาด** มีการกำหนดช่วงของขนาดที่มีความน่าจะเป็น ดวงตา ปาก และจมูก มากที่สุด ถ้าขนาดใหญ่เกินกว่าขอบเขตที่กำหนด ก็สามารถระบุได้ว่าเป็นตำแหน่งของส่วนบริเวณผม
- **ตำแหน่ง** จากสมมติฐานที่ว่ารายละเอียดบนใบหน้าจะอยู่บริเวณกลางของวงรี จึงใช้เป็นตัวบ่งชี้ในการกำหนดพื้นที่ความน่าจะเป็นที่มีโอกาสตรวจพบ

สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย




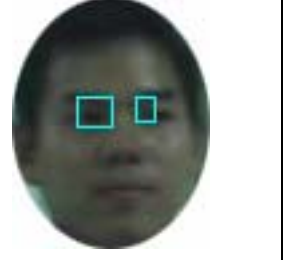





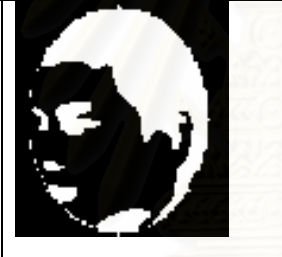

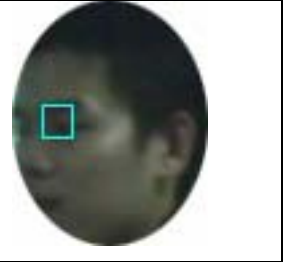


รูปที่ 7 กระบวนการหารายละเอียดของใบหน้า

### 3.1.2.2.1 ผลการทดลองและวิเคราะห์






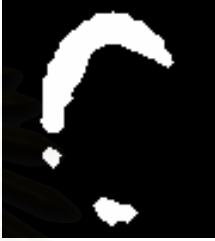
ตารางที่ 3 แสดงถึงตัวอย่างผลการทดลองในการตรวจหาดวงตาของภาพใบหน้าบุคคล โดยระบบจะเลือกบริเวณที่มีความน่าจะเป็นมากที่สุด ซึ่งเหตุผลที่นำกระบวนการ morphological มาประยุกต์ใช้ก็เพื่อลดผลกระทบของบริเวณที่เป็นสัญญาณรบกวน เพื่อให้เหลือเฉพาะบริเวณที่สำคัญ จะเห็นว่าจากผลการทดลองสามารถหาตำแหน่งดวงตาได้ แต่รายละเอียดอื่น ๆ ก็ถูกตัดทิ้งไปด้วย ซึ่งเป็นปัญหาที่เกิดขึ้นในการทดลองเบื้องต้นนี้ และปัญหาอีกประการหนึ่งก็คือ เรื่องของแสงหรือค่าความสว่างของภาพใบหน้า ซึ่งส่งผลโดยตรงถ้าความสว่างไม่เพียงพอ ก็จะจำแนกส่วนบริเวณที่เป็นรายละเอียดของใบหน้าออกมาไม่ได้ ดังตารางที่ 4 แถวล่างและด้านศีรษะของบุคคลที่ตรวจหาก็ส่งผลกระทบบังตารางที่ 4 แถวบน ซึ่งส่วนบริเวณผมของบุคคลเข้ามาใกล้กับบริเวณดวงตามากเกินไป

ตารางที่ 3 ผลการตรวจหารายละเอียดใบหน้า

ภาพใบหน้า	Thresholding	Morphological	Facial feature
			
			
			

สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4 ผลการตรวจหารายละเอียดใบหน้าที่ถูกผิดพลาด

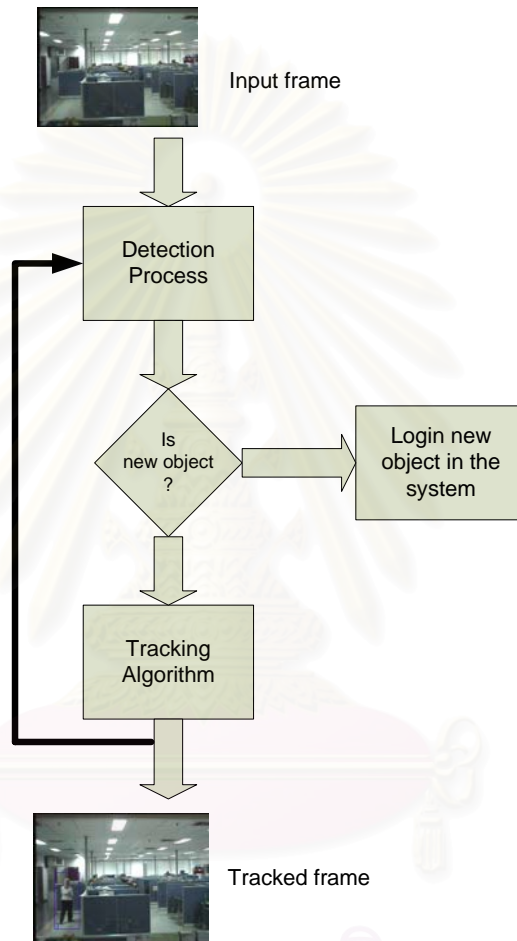
ภาพใบหน้า	Thresholding	Morphological
		
		

### 3.2. การทำงานร่วมกันสำหรับการตรวจหาติดตามใบหน้าอย่างต่อเนื่องระหว่างกล้องวีดีทัศน์ 2 ตัว

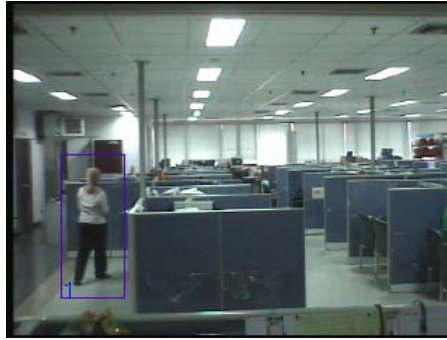
ปัญหาที่เกิดขึ้นสำหรับการทำงานร่วมกันในลักษณะนี้ คือ ความไม่มีประสิทธิภาพในการติดตามบุคคล โดยในระบบเดิมสามารถรองรับการติดตามบุคคลได้เพียง 1 คนเท่านั้น ภายในมุมมองการรับภาพ ดังนั้นงานวิจัยในส่วนงานนี้ จึงมุ่งเน้นในการพัฒนาและปรับปรุงการติดตามบุคคลภายในมุมมองการรับภาพของกล้องให้มีประสิทธิภาพในการติดตามมากขึ้น รวมทั้งสามารถติดตามบุคคลได้ในเวลาเดียวกันมากกว่า 1 คน

รูปแบบโดยรวมของระบบแสดงได้ดังรูปที่ 8 โดยในที่นี้จะทดลองกับกล้องอยู่กับที่ ดังนั้นขั้นตอนวิธีที่เลือกใช้สำหรับการตรวจหาบุคคล จะเลือกใช้กระบวนการลบฉากหลัง โดยจะมีการเปลี่ยนรูปแบบของกระบวนการตรวจหา คือ เดิมในกระบวนการตรวจหาจะประมวลผลทั้งภาพขาเข้า ซึ่งส่งผลต่อเวลาในการประมวลผล ดังนั้นจึงมีการเปลี่ยนรูปแบบ โดยแบ่งออกเป็น 2 ส่วน คือ ส่วนแรก จะตรวจหาเฉพาะบริเวณขอบของภาพ (ในการทดลองเลือกใช้ระยะ 20 จุดภาพจากขอบภาพทุกด้าน) เป้าหมายของรูปแบบส่วนแรกนี้ คือ การตรวจหาบุคคลใหม่ที่จะเข้ามาในระบบ สำหรับส่วนที่สอง จะตรวจหาเฉพาะบริเวณ คือ ในกรณีที่ระบบได้มีการลงทะเบียนบุคคลที่ตรวจหาเข้ามาในระบบแล้ว ระบบจะทำนายตำแหน่งถัดไปของบุคคลนี้ รวมทั้งกำหนดขอบเขตบริเวณของการตรวจหา (Window search) ขอบเขตบริเวณนี้จะเป็นบริเวณตรวจหาสำหรับเฟรมภาพถัดไปเฉพาะของบุคคลนี้ ดังนั้นจะไม่มีกรตรวจหาในตำแหน่งที่ไม่มีความจำเป็น ซึ่งจะช่วยลดเวลาในการ

ประมวลผลได้อย่างมาก ตัวอย่างผลของการตรวจหาแสดงได้ดังรูปที่ 9 ระบบจะมีการบอกถึงลำดับหรือตัวเลขกำกับของแต่ละบุคคลที่เคลื่อนที่ภายในบริเวณมุมมองการรับภาพ แต่ในระบบนี้จะไม่มีการจดจำรูปลักษณะของบุคคล กล่าวคือ เมื่อบุคคลนั้นเคลื่อนออกจากมุมมองการรับภาพแล้ว กลับเข้ามาใหม่ ระบบก็จะถือว่าเป็นคนละบุคคลกัน



รูปที่ 8 ไคอะแกรมการทำงานของกระบวนการติดตามบุคคล



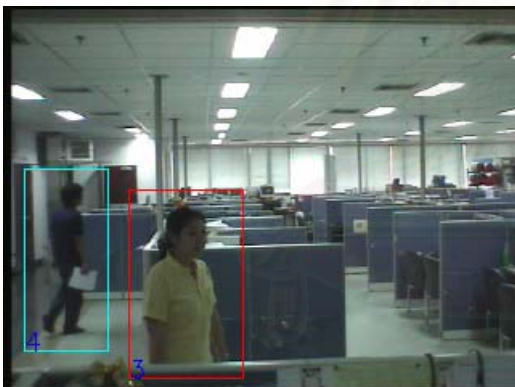
รูปที่ 9 ตัวอย่างผลของการตรวจหาและติดตามบุคคล



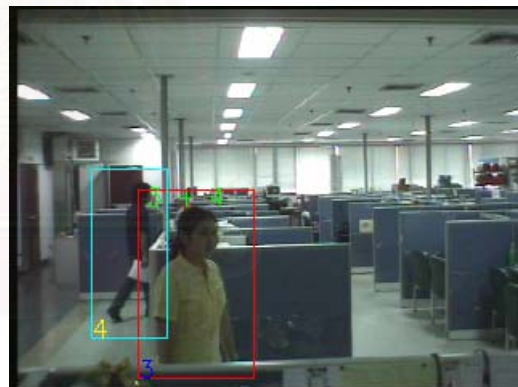
(ก)



(ข)



(ค)



(ง)

รูปที่ 10 ผลการตรวจหาและติดตามในกรณีมีบุคคลมากกว่า 1 คน และกรณีมีการซ้อนทับกันระหว่างบุคคล

ผลการตรวจหาและติดตามในกรณีมีบุคคลมากกว่า 1 คน แสดงได้ดังรูปที่ 10 (ค) จะเห็นว่ามีการกำหนดเลขกำกับเป็นบุคคลที่ตรวจหาได้ในลำดับที่ 3 และ 4 และในกรณีที่มีการซ้อน



กันระหว่างบุคคลภายในภาพ ระบบจะมีตัวเลขกำกับว่า ณ ขณะนั้นมีการซ้อนกันระหว่างบุคคลใด กับบุคคลใด

### วิเคราะห์และแนวทางแก้ไขปัญหา

ปัญหาที่เกิดขึ้นของระบบ คือ ในกรณีที่มีการเหลื่อมกันหรือซ้อนทับกัน ระบบยังมีการจำแนกบุคคลที่ผิดพลาด ซึ่งจำเป็นต้องมีขั้นตอนวิธีในการจำแนกบุคคลกรณีก่อนการซ้อนทับ และหลังการซ้อนทับ โดยต้องนำลักษณะสำคัญของบุคคลมาเป็นตัวบ่งชี้ เช่น ลักษณะทางด้านสีของบุคคล สีของเสื้อ สีผิว ลักษณะรูปร่างและขนาด ลักษณะทิศทางของการเคลื่อนที่ เป็นต้น และปัญหาอีกประการหนึ่ง คือ ในกรณีการลงทะเบียนวัตถุที่ตรวจหาได้เข้ามาในระบบ จำเป็นต้องมีกระบวนการเพื่อใช้ในการจำแนกว่าวัตถุที่ตรวจหาได้เป็นบุคคลหรือว่าสิ่งของอื่น ๆ



สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย

## ภาคผนวก ข

### 1. จุดประสงค์ของงานวิจัย

1.1. ศึกษาทฤษฎีและหลักการพื้นฐานของการทำงานร่วมกันของระบบกล้องวิดีโอทัศน์ และศึกษาทฤษฎี หลักการพื้นฐานของระบบติดตามบุคคลโดยใช้กล้องวิดีโอทัศน์ในรูปแบบต่างๆ จากบทความงานวิจัยที่ผ่านมา

1.2. ศึกษาและทำการทดลองตัวอย่างระบบการตรวจหาและติดตามบุคคลด้วยการใช้กล้องวิดีโอทัศน์ประเภท PTZ สองตัวให้ทำงานร่วมกัน

1.3. ศึกษาการทำงานร่วมกันของระบบกล้องมากกว่าสองตัวและสถานการณ์ที่เหมาะสมต่อการทำงานของระบบ

### 2. รายละเอียดโดยย่อของงานวิจัย

#### 2.1. งานวิจัยที่เกี่ยวข้องกับระบบการทำงานของกล้องวิดีโอทัศน์หลายตัว

งานวิจัยที่เกี่ยวข้องกับการทำงานของระบบกล้องวิดีโอทัศน์หลายตัว ที่เคยเผยแพร่เป็นบทความทางวิชาการนั้น ส่วนใหญ่จะนำระบบกล้องหลายตัวนี้ มาใช้งานเป็นระบบตรวจตราและสังเกตการณ์ต่างๆ (Surveillance System) สรุปงานวิจัยที่เกี่ยวข้องกับระบบกล้องหลายตัวได้ดังนี้

J.Orwell, P.Remagnino and G.A.Jones นำเสนอการทำงานของระบบกล้องหลายตัวแบบอยู่กับที่ทุกตัว (Stationary Camera) และใช้เทคนิคของการลบฉากหลัง (Background Subtraction) โดยสมมุติให้แบบฉากหลัง มีรูปแบบการกระจายเป็นแบบ Mixture of Gaussian และใช้ Cross entropy distance measure ข้อดีของงานวิจัยนี้คือ ใช้หลักของความน่าจะเป็นและการวัดระยะทางอย่างง่าย ทำให้ระบบที่ออกแบบมีความรวดเร็วในการประมวลผลสูง แต่ไม่รองรับสถานการณ์ที่เปลี่ยนแปลงของฉากหลังได้

T. Ellis งานวิจัยนี้นำเสนอการทำงานของระบบกล้องหลายตัว ที่สามารถใช้งานได้ภายนอกอาคารเพื่อทำหน้าที่ตรวจหาและติดตามการเคลื่อนที่ของคนเดินถนนและยานพาหนะ โดยออกแบบระบบให้สามารถปรับตัวให้เข้ากับปัจจัยทางด้านความส่องสว่าง (Illumination) ภายนอกอาคาร

S. Denman, C. Fookes, J. Cook and C. Davoren, ระบบกล้องหลายตัวในงานวิจัยนี้ได้ใช้กล้องจำนวน 4 ตัว นำมาใช้ตรวจหาและติดตามยานพาหนะ บริเวณลานจอดรถ ซึ่งระบบจะมีความสามารถในการตรวจหาและติดตามในเวลาจริงได้ และใช้เทคนิคของ Motion Detection และ Optical Flow

F. Porikli, A. Divakaran นำเสนอวิธีการเทียบมาตรฐานกล้องวิดีโอทัศน์เพื่อใช้ในระบบการตรวจหาและติดตามบุคคลด้วยกล้องวิดีโอทัศน์ในระบบมากกว่าสองตัว งานวิจัยนี้ได้ให้ความสำคัญกับการสร้างให้ระบบกล้องวิดีโอทัศน์ทำงานร่วมกันอย่างอัตโนมัติ มีการเปรียบเทียบมาตรฐานกล้องวิดีโอทัศน์อย่างถูกต้องแม่นยำ และสามารถตรวจจับและติดตามบุคคลหรือวัตถุที่สนใจได้อย่างถูกต้อง สิ่งในงานวิจัยนี้ให้ความสำคัญมากที่สุด คือ การมองถึงปัญหาการเทียบมาตรฐานสีของกล้องวิดีโอทัศน์ จากงานวิจัยที่ผ่านมา จะกำหนดให้แสงสว่างบริเวณที่กล้องแต่ละตัวติดตั้งนั้น มีค่าเท่ากันหรือประมาณให้มีค่าใกล้เคียงกัน ซึ่งเป็นสิ่งที่ทำให้ผลการตรวจหาและติดตามบุคคลที่สนใจมีความผิดพลาด แต่ในความเป็นจริงแล้วปัญหาความสว่างของบริเวณที่ติดตั้งกล้องวิดีโอทัศน์แต่ละตัวที่ต่างกันนั้น ไม่สามารถควบคุมได้

การทำงานของระบบกล้องวิดีโอทัศน์นั้น จะทำงานภายใต้สถานการณ์ที่มีข้อจำกัดทางด้านแสงสว่างที่ต่างกันในแต่ละบริเวณ และการติดตั้งกล้องแต่ละตัวจะไม่มีตำแหน่งที่เหลื่อมกัน (Non Overlapping Multi Cameras System)

A. Gilbert and R. Bowden งานวิจัยนี้นำเสนอการติดตามบุคคลอัตโนมัติระหว่างกล้องวิดีโอทัศน์ (Inter-Camera Tracking) โดยหลักการจะอาศัยการติดตามบุคคลที่ผ่านเข้ามาในระบบกล้องวิดีโอทัศน์แต่ละตัวอย่างต่อเนื่อง (Handover) นั่นคือ จะให้กล้องตัวหนึ่งส่งข้อมูลบุคคลที่มีการเคลื่อนที่ให้กับกล้องตัวอื่นๆ เพื่อทำหน้าที่ติดตามบุคคลอย่างต่อเนื่อง จากงานวิจัยนี้ได้ให้ความสำคัญกับ ความสัมพันธ์ระหว่างกล้องวิดีโอทัศน์แต่ละตัวโดยอาศัยหลักการวิธีเทียบมาตรฐานทางสี (Color Calibration) ของกล้องวิดีโอทัศน์ ใช้งานร่วมกับการติดตามบุคคลด้วยระบบกล้องวิดีโอทัศน์หลายตัวด้วย

## 2.2. ปัจจัยที่มีผลต่อการทำงานของระบบกล้องหลายตัว

จากการศึกษาบทความงานวิจัยเกี่ยวกับระบบกล้องวิดีโอทัศน์เพื่อการตรวจหาและติดตามบุคคล สามารถสรุปปัจจัยทางสภาพแวดล้อมและลักษณะการติดตั้งมุมมองของกล้องวิดีโอทัศน์ ที่มีผลต่อการทำงานของระบบ ได้ดังนี้

### 2.2.1. ปัจจัยทางสภาพสิ่งแวดล้อม (Environment Conditions)

- ปัจจัยทางแสง (Light condition) บริเวณที่กล้องแต่ละตัวติดตั้งมี ความสว่างเท่ากัน (Uniform lighting condition)
- บริเวณที่กล้องแต่ละตัวติดตั้งมีความสว่างเท่ากัน (Under different lighting condition)
- ปัจจัยทางรังสี (Radiometric condition)
- ปัจจัยทางความผิดเพี้ยนของเลนส์ (Lens Distortion condition)

### 2.2.2. ลักษณะการติดตั้งมุมมองของระบบกล้อง (Multi Cameras View Setup)

- การติดตั้งระบบกล้องที่มีการเหลื่อมกัน (Overlapping cameras view)
- การติดตั้งระบบกล้องที่ไม่มีการเหลื่อมกัน (Non-Overlapping cameras view)

## 2.3. ระเบียบวิธีของระบบติดตามบุคคลด้วยระบบกล้องวิดีโอทัศน์

จากการศึกษาบทความงานวิจัยเกี่ยวกับระบบกล้องวิดีโอทัศน์เพื่อการตรวจหาและติดตามบุคคล สามารถสรุปปัจจัยทางสภาพแวดล้อมและลักษณะการติดตั้งมุมมองของกล้องวิดีโอทัศน์ ที่มีผลต่อการทำงานของระบบ ได้ดังนี้

### 2.3.1. ขั้นตอนการเตรียมข้อมูลและการตรวจหาข้อมูล (Pre-Processing)

- การตรวจจับและแยกส่วน (Detection and Segmentation) ประกอบไปด้วย

- วิธีการลบฉากหลัง (Background Subtraction or Motion Detection)
- วิธีการตรวจจับพื้นที่ที่มีการเคลื่อนที่ (Moving Region)
- วิธีการตรวจจับทางสี (Color Detection)
- วิธีการสร้างโมเดลจำลองของมนุษย์ (Using Model of Human Head Appearance)

### 2.3.2. การดึงคุณลักษณะสำคัญที่ต้องการ (Feature Extraction)

### 2.3.3. ขั้นตอนการติดตามบุคคล (People Tracking Method)

- Image based method ได้แก่
  - วิธี Template Matching Method เช่น สี พื้นที่ หรือ ส่วนของร่างกาย เป็นต้น
- Trajectory based method ได้แก่
  - วิธีของ Kalman Filter (KF) , วิธี Optical flow, Nearest Neighbor Method (NNM) Probabilistic, Data Association filter (PDAF), Joint Visual Probabilistic Data Association filter (JVPDAF) เป็นต้น

## 2.4. ตัวอย่างระบบการติดตามใบหน้าและการเคลื่อนที่ของบุคคลกล้องวิดีโอประเภท PTZ 2 ตัว

### 2.4.1 ขั้นตอนวิธีของระบบการติดตามใบหน้าและการเคลื่อนที่ของบุคคลด้วยกล้องวิดีโอ 2 ตัว

ขั้นตอนวิธีในการตรวจหาและติดตามบุคคลด้วยกล้องวิดีโอ PTZ ที่ได้ศึกษานี้ มีแนวความคิดหลักของการตรวจหาและติดตามบุคคลโดยจะอาศัยการตรวจจับส่วนของผิวบริเวณใบหน้าและการเคลื่อนที่ของบุคคลภายในเฟรมวิดีโอ ดังนั้นขั้นตอนของการติดตามใบหน้าและการเคลื่อนที่ของบุคคลด้วยกล้องวิดีโอ 2 ตัวนี้ จะประกอบไปด้วย

- ขั้นตอนส่วนการตรวจหาและติดตามใบหน้า
- ขั้นตอนส่วนควบคุมการส่ายและกัมเมยของกล้องวิดีโอ PTZ ติดตามบุคคล เป้าหมาย
- ขั้นตอนส่วนการทำงานร่วมกันระหว่างกล้อง 2 ตัวเพื่อติดตามใบหน้าบุคคลเป้าหมาย

#### 2.4.1.1 ขั้นตอนการตรวจหาและติดตามใบหน้า

ตัวอย่างระบบการตรวจหาและติดตามใบหน้าที่ได้ศึกษานี้ จะอาศัยการทำงานร่วมกันของสองวิธีของการตรวจหาให้ทำงานร่วมกันดังนี้

การตรวจหาด้วยสีผิว ปัญหาที่เกิดขึ้นในการตรวจหาส่วนสีผิว คือ ปัญหาเกี่ยวกับสิ่งแวดล้อม ฉากหลังที่มีสีที่คล้ายกับส่วนสีผิว (ส่วนมากวัตถุที่มีผลกระทบจะมีองค์ประกอบสีหลักเป็นสีแดง) ซึ่งทำให้เกิดความยุ่งยากในการวิเคราะห์บริเวณ จึงมีการประยุกต์ใช้การตรวจหาด้วยลักษณะการเคลื่อนที่ร่วมด้วย ซึ่งจะสามารถลดปัญหาความผิดพลาดในเรื่องสีผิวกับสีของสิ่งแวดล้อมได้

กระบวนการตรวจหาการเคลื่อนที่ กระบวนการเบื้องต้นของการตรวจหาการเคลื่อนที่โดยทั่วไป คือ การหาความแตกต่างระหว่างเฟรม (Frame Differencing) ผลที่ได้จะเป็นบริเวณขอบของการเคลื่อนที่และบริเวณบางส่วนภายในขอบเขตการเคลื่อนที่ ที่มีค่าความต่างของความเข้มสีอย่างชัดเจน ซึ่งผลที่ได้จะไม่ได้เป็นโครงร่างโดยรวมของบุคคล ทำให้มีความยุ่งยากในขั้นตอนการวิเคราะห์ผลการตรวจหา เนื่องจากการตรวจหาสีผิว จะเห็นว่าปัญหาหลักที่เกิดขึ้น คือ อาจจะต้ององค์ประกอบของสีที่ใกล้เคียงกับสีผิว ดังนั้นคุณสมบัติของการเคลื่อนที่จึงได้ถูกนำมาประยุกต์ใช้ เพื่อใช้แยกบริเวณส่วนที่มีการเคลื่อนที่เท่านั้นออกจากฉากหลัง ซึ่งจะตัดผลกระทบที่เกิดจากการตรวจหาสีผิวได้

#### 2.4.1.2 ขั้นตอนควบคุมการส่ายกัมเมยและซูมของกล้องวิดีโอ PTZ ติดตามใบหน้าบุคคลเป้าหมาย

การควบคุมการส่ายกัมเมย และซูมของกล้องวิดีโอ PTZ เพื่อติดตามใบหน้าบุคคลเป้าหมายนั้น เป็นกระบวนการสำหรับคำนวณหามุมและอัตราการซูม ในการเปลี่ยนมุมมองของกล้องวิดีโอ โดยมีเป้าหมาย คือ พยายามเปลี่ยนมุมมองการรับภาพของกล้องให้ใบหน้า

บุคคลเป้าหมายมาอยู่บริเวณกึ่งกลางภาพ ตัวอย่างการปรับเปลี่ยนมุมมองของกล้องวิดีโอให้บุคคลเป้าหมายอยู่บริเวณกึ่งกลางของภาพ

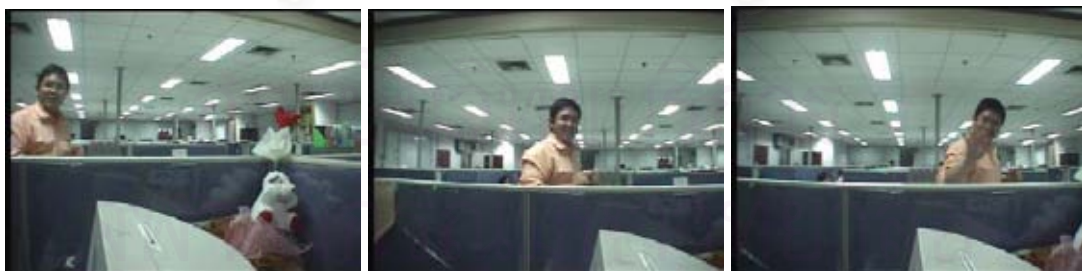
ในกรณีที่ใบหน้าเป้าหมายอยู่บริเวณใกล้กับกึ่งกลางภาพอยู่แล้วจากระบบจะไม่จำเป็นต้องสั่งให้กล้องวิดีโอปรับเปลี่ยนมุมมอง ดังนั้นจึงได้กำหนดช่วงขอบเขตในภาพเพื่อใช้เป็นขอบเขตในการเคลื่อนที่ ถ้าตรวจหาได้ว่าใบหน้ายังอยู่ในช่วงที่กำหนดที่ไม่ต้องมีการปรับเปลี่ยนมุมมอง (Non-moving boundary) ก็จะไม่สั่งกล้องให้เคลื่อนที่ จนกว่าใบหน้าเป้าหมายจะอยู่นอกขอบเขตก็จะสั่งกล้องให้เคลื่อนที่

#### 2.4.1.3 การทำงานร่วมกันระหว่างกล้องวิดีโอ PTZ สองตัวเพื่อติดตามใบหน้าบุคคลเป้าหมาย

การทำงานร่วมกันระหว่างกล้องวิดีโอ PTZ 2 ตัว เช่น ในสถานการณ์เมื่อกล้องวิดีโอตัวแรกสามารถตรวจหาพบบุคคล ก็จะเก็บข้อมูลลักษณะ (Feature Data) ของบุคคลนี้ไว้ซึ่งข้อมูลชุดนี้จะถูกส่งไปให้กับกล้องวิดีโอตัวที่สอง กล้องวิดีโอตัวที่สองนี้จะนำข้อมูลนี้เพื่อใช้เปรียบเทียบบุคคลที่เข้ามาในมุมมองกล้องว่าเป็นบุคคลเดียวกันหรือไม่ ถ้าใช่ก็จะทำการตรวจหาและทำการถ่าย ก้มเงย และซูม กล้องวิดีโอเพื่อติดตามบุคคลต่อเนื่อง เพื่อสามารถจัดเก็บภาพบุคคลที่สนใจได้อย่างต่อเนื่องจากกล้องวิดีโอตัวหนึ่งไปยังอีกตัวหนึ่ง

#### 2.4.1. การทดลองและผลการทดลอง

##### 2.4.1.1. การทดลองที่ 1 ทดสอบการควบคุมการถ่ายก้มเงยซูมติดตามบุคคลของกล้องวิดีโอ



(ก)

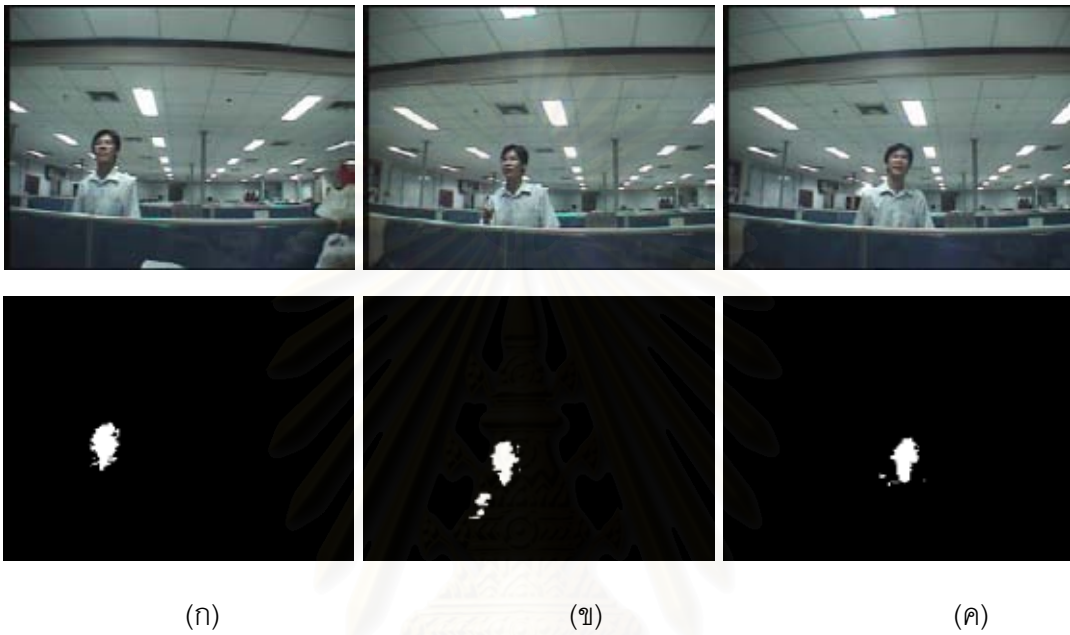
(ข)

(ค)

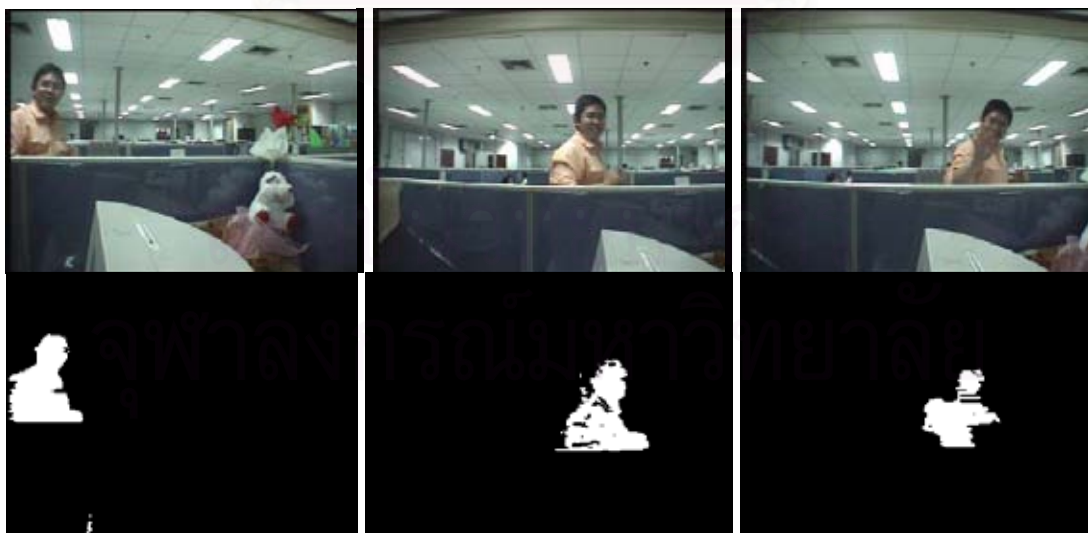
รูปที่ 2.4.1 แสดงการก้มเงยถ่ายเพื่อหาบุคคลเป้าหมายและพยายามให้บุคคลอยู่บริเวณกึ่งกลางภาพ

จากรูปที่ 2.4.1. (ก) เมื่อบุคคลมีการเคลื่อนที่ใกล้กล้องวิดีโอที่มุมด้านซ้ายของกล้อง กล้องวิดีโอแบบ PTZ จะทำการปรับภาพให้บุคคลที่มีการเคลื่อนที่เข้ามาภายในกล้อง เลื่อนตำแหน่งมาอยู่ ณ ตำแหน่งกึ่งกลางภาพ ดังรูปที่ 2.4.1. (ข) และเมื่อบุคคลเคลื่อนที่ต่อไปยังมุมขวาของกล้อง จะเห็นได้ว่ากล้องจะพยายามปรับให้บุคคลนั้น อยู่ตรงกึ่งกลางของกล้องเช่นเดียวกัน

#### 2.4.1.2. การทดลองที่ 2 ทดสอบการการตรวจหาและติดตามด้วยสีผิว



รูปที่ 2.4.2. แสดงการตรวจหาและติดตามด้วยสีผิว



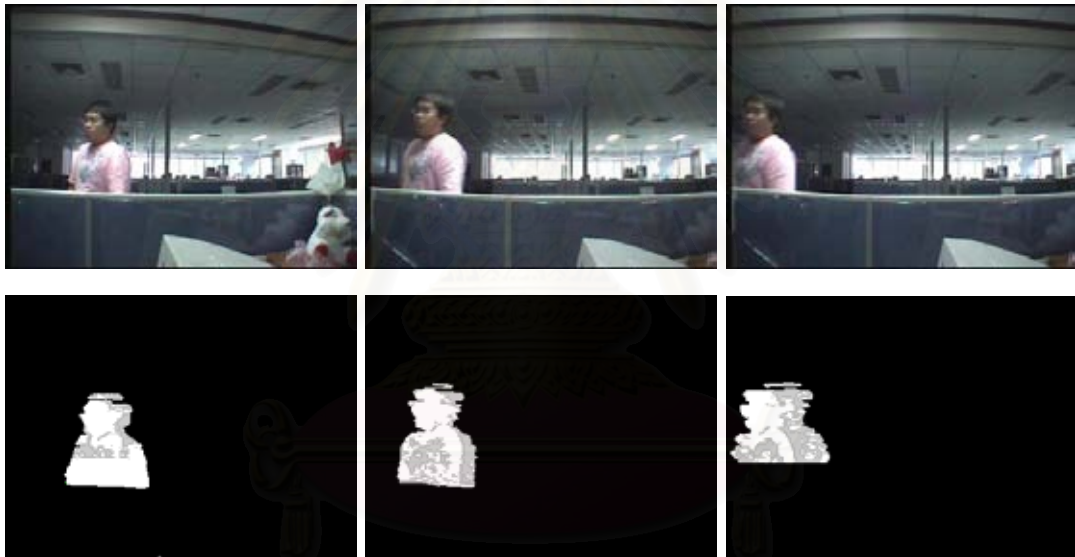
รูปที่ 2.4.3. แสดงการตรวจหาและติดตามด้วยสีผิวกรณีที่สิ่งแวดล้อม ที่มีองค์ประกอบทางสีเหมือนสีผิว เช่น สีเสื้อมีสีเดียวกับสีผิว



จากรูปที่ 2.4.2 แสดงการตรวจหาและติดตามบุคคลโดยอาศัยสีผิว ได้แก่ บริเวณใบหน้า ดังแสดงในรูป 4.2.1(ก) และ(ค) และสามารถตรวจหาส่วนอื่นของร่างกายได้อีก เช่น แขน ดังรูปที่ 4.2.1(ข)

จากรูปที่ 2.4.3 แสดงการตรวจหาและติดตามบุคคล โดยอาศัยสีผิวในกรณีที่สิ่งแวดล้อมมีองค์ประกอบของสีเหมือนกับสีผิว (ในกรณีนี้คือสีเขียว) ทำให้ระบบนี้สามารถตรวจหาส่วนที่เป็นสีเขียวได้ด้วย

### การทดลองที่ 3 ทดสอบการการตรวจหาและติดตามด้วยการเคลื่อนที่



(ก) เฟรมที่  $t$

(ข) เฟรมที่  $t+1$

(ค) เฟรมที่  $t+2$

### รูปที่ 2.4.5. แสดงการตรวจหาและติดตามด้วยการเคลื่อนที่

จากรูปที่ 2.4.5. แสดงการตรวจหาและติดตามด้วยการเคลื่อนที่ โดยส่วนที่ระบุเป็นสีเทาจะเป็นส่วนที่ระบุว่าวัตถุในเฟรมวิดีโอที่ติดกันนั้น มีการเคลื่อนที่ ในเฟรมของวิดีโอที่ติดกัน พบว่า ส่วนที่เป็นสีเทาคือส่วนที่ระบบสามารถตรวจจับได้ว่า บุคคลภายในเฟรมวิดีโอที่มีการเคลื่อนที่

2.4.1.4. การทดลองที่ 4 ทดสอบการทำงานร่วมกันเพื่อระบุบุคคลต่อเนื่องระหว่างกล้องวิดีโอทัศน์ประเภท PTZ สองตัวแบบ Non-Overlapped ในสถานการณ์ต่างๆ

ตารางที่ 9.4. การทำงานร่วมกันเพื่อระบุบุคคลต่อเนื่องระหว่างกล้องวิดีโอทัศนัประภท PTZ สองตัวแบบไม่เหลื่อมกัน

	มุมมองกล้องวิดีโอทัศนั PTZ ตัวที่ 1	มุมมองกล้องวิดีโอทัศนั PTZ ตัวที่ 2
บุคคลเป้าหมายเดินเข้า มุมมองกล้องวิดีโอทัศนั PTZ ตัว ที่ 1		---
กล้องวิดีโอทัศนัตัวที่ 1 สาย กัม เงยและซูมติดตามเป้าหมาย และเก็บข้อมูลลักษณะสีและ ข้อมูลการเคลื่อนที่ไปทางขวา		---
กล้องวิดีโอทัศนัตัวที่ 1 สาย กัม เงยและซูมติดตามเป้าหมาย โดยจะปรับมุมมองให้บุคคล อยู่ในตำแหน่งตรงกลางของ ภาพ		---
บุคคลเป้าหมายเคลื่อนเข้า มุมมองกล้องวิดีโอทัศนัที่ 2 และ เริ่มทำการตรวจสอบข้อมูล ลักษณะสีและข้อมูลทิศ	---	

<p>ทางการเคลื่อนที่</p>		
<p>การเคลื่อนที่เป็นไปในทิศทางขวา และปรับมุมมองให้บุคคลอยู่ในตำแหน่งตรงกลางของภาพ</p>	<p>---</p>	
<p>กล้องวีดิทัศน์ตัวที่ 2 ยังเปลี่ยนมุมมองติดตาม ถ้าบุคคลเป้าหมายยังอยู่ในอาณาบริเวณที่สามารถถ่ายและก้มเงยไปได้</p>	<p>----</p>	

#### 2.4.2. วิเคราะห์และสรุปผลการทดลอง

จากการทดสอบระบบดังกล่าว พบว่า การตรวจหาสีผิวนั้นมีข้อจำกัดเมื่อสิ่งแวดล้อมภายในวิดิทัศน์ มีองค์ประกอบของสีที่ใกล้เคียงหรือเหมือนกับสีผิวมนุษย์ ซึ่งจะทำให้ผลการตรวจหาและติดตามบุคคลมีความผิดพลาด ระบบตัวอย่างที่ได้ศึกษา จึงเพิ่มส่วนของการตรวจจับและติดตามบุคคลด้วยการเคลื่อนที่ ซึ่งผลการทดลองการตรวจหาและติดตามจะมีประสิทธิภาพเพิ่มขึ้นด้วย

## 2.5 การติดตามใบหน้าและการเคลื่อนที่ของบุคคลโดยใช้กล้องวิดีโอทัศนมากกว่า 2 ตัว

2.5.1 ที่มาและปัญหาของระบบการติดตามใบหน้าและการเคลื่อนที่ของบุคคลโดยใช้กล้องวิดีโอทัศนมากกว่า 2 ตัว

จากการศึกษาและทดลองระบบการติดตามใบหน้าและการเคลื่อนที่ของบุคคลโดยใช้กล้องวิดีโอทัศน 2 ตัวที่ผ่านมา พบว่าขั้นตอนที่ต้องให้ความสำคัญ คือ ขั้นตอนการประมวลผลภาพ ซึ่งข้อมูลที่ได้รับจากกล้องแต่ละตัวนั้น ตัวข้อมูลจำเป็นต้องมีความถูกต้อง แม่นยำและเชื่อถือได้ เนื่องจากระบบจะมีการใช้ข้อมูลร่วมกันมาประมวลผลระหว่างกล้อง และเนื่องจากความต้องการข้อมูลที่ถูกต้องนั่นเอง ปัญหาของระบบการติดตามใบหน้าและการเคลื่อนที่ของบุคคลโดยใช้กล้องวิดีโอทัศน 2 ตัวที่ได้ศึกษามานั้น จึงเกิดปัญหาบางประการเกี่ยวกับการทำงานของระบบได้แก่

- กรณีที่สิ่งแวดล้อมในการติดตั้งวิดีโอทัศนทั้ง 2 ตัว มีองค์ประกอบของสีที่ใกล้เคียงหรือเหมือนกับสีผิว มนุษย์ ซึ่งจะทำให้ผลการตรวจหาและติดตามบุคคลมีความผิดพลาด
- การทำงานของระบบจะทำงานได้ดีมีประสิทธิภาพเมื่อมีการติดตั้งกล้องวิดีโอทัศนทั้ง 2 ตัวในบริเวณที่มีสภาพแวดล้อมที่ใกล้เคียงกัน เช่น ความสว่างที่มีปริมาณเท่ากัน หรือไม่แตกต่างกันมากนัก แต่ในการใช้งานเพื่อประโยชน์อย่างแท้จริงแล้ว ระบบควรรองรับการทำงานเมื่อมีการติดตั้งกล้องวิดีโอทัศนที่ต่างสถานที่กัน เพราะฉะนั้นจึงจำเป็นต้องให้ความสำคัญกับปัจจัยอื่น เมื่อมีการติดตั้งกล้องวิดีโอทัศนที่ต่างบริเวณกันด้วย

เนื่องจากระบบการติดตามใบหน้าและการเคลื่อนที่ของบุคคลด้วยกล้องวิดีโอทัศน 2 ตัว นั้น สามารถทำงานได้ดีในสถานการณ์ที่กล้องทั้งสองตัวอยู่ในสภาพแวดล้อมเดียวกันดังกล่าว ปัญหาและข้อจำกัดต่างๆ ที่เกิดขึ้น จะสามารถแก้ไขด้วยการออกแบบระบบการทำงานร่วมกันระหว่างกล้องวิดีโอทัศนที่มากกว่าสองตัวและอาศัยการเทียบมาตรฐานของกล้องวิดีโอทัศนร่วมด้วย

ตัวอย่างงานวิจัยเกี่ยวกับระบบการติดตามบุคคลด้วยกล้องวิดีโอทัศนมากกว่าสองตัว เช่น งานวิจัยของ G. Unal<sup>1</sup>, A. Yezzi, H. Baker ได้นำเสนอปัญหาของการเทียบมาตรฐานพารามิเตอร์ของกล้องวิดีโอทัศน ด้วยวิธีการเทียบมาตรฐานกล้องวิดีโอทัศนด้วยสีด้วย ซึ่งหลักการ คือนำภาพที่ได้จากกล้องวิดีโอทัศนแต่ละตัว มาคำนวณหาค่าสีภายในภาพของกล้องแต่ละตัว แล้วเปรียบเทียบกับวัตถุที่ใช้ในการเปรียบเทียบมาตรฐานทางสี โดยมีจุดประสงค์เพื่อให้ระบบกล้องโครงการวิจัยร่วมฯ ปีงบประมาณ 2549 556 จัดทำเมื่อ 10 มกราคม 2551

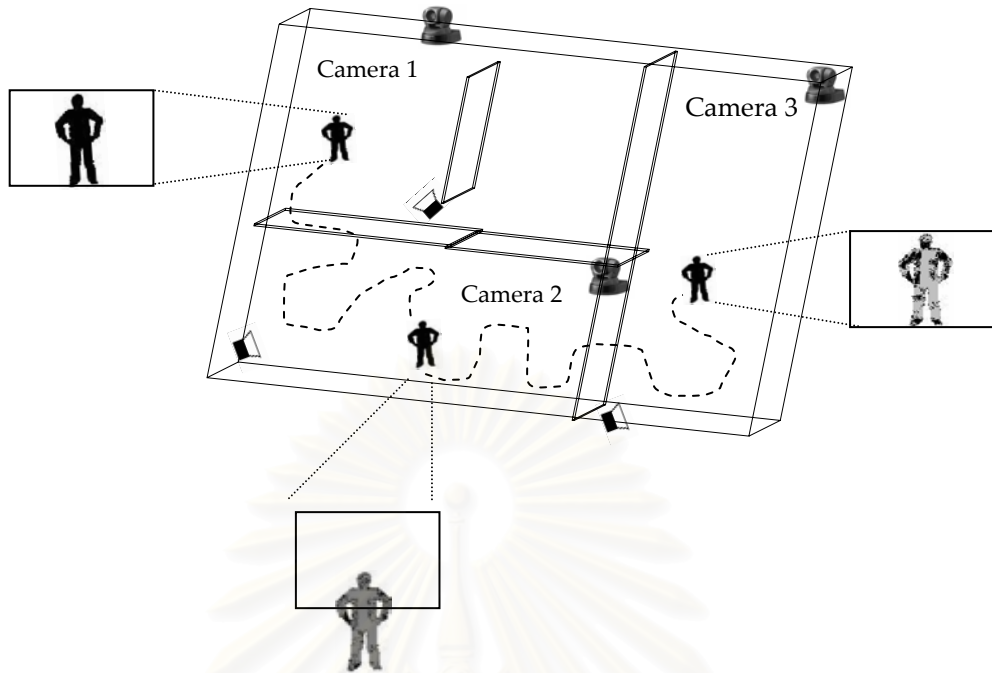
วิธีที่คนหลายตัวนี้ ทำงานเปรียบเทียบได้ว่าเป็นกล่องวิทัศน์ตัวเดียวกัน มองภาพและประมวลผลทางสีของภาพได้อย่างถูกต้องแม่นยำมากยิ่งขึ้น

สิ่งที่งานวิจัยนี้ให้ความสำคัญมากที่สุด คือ การมองถึงปัญหาการเทียบมาตรฐานสีของกล่องวิทัศน์ ซึ่งจากงานวิจัยในระยะแรกที่ผ่านมา จะตั้งอยู่ในสมมุติฐานว่า ความส่องสว่างบริเวณที่กล่องแต่ละตัวที่ติดตั้งนั้น มีค่าเท่ากัน หรือประมาณให้มีค่าใกล้เคียงกัน ซึ่งเป็นสิ่งที่ทำให้ผลการตรวจหาและติดตามบุคคลที่สนใจมีความผิดพลาด แต่ในความเป็นจริงแล้วปัญหาความสว่างของบริเวณที่ติดตั้งกล่องวิทัศน์แต่ละตัวที่ต่างกันนั้น ไม่สามารถควบคุมได้

สรุปได้ว่า ปัญหาของระบบการทำงานร่วมกันระหว่างกล่องวิทัศน์ที่มากกว่าสองตัวที่สำคัญนั้น คือ ขั้นตอนของการประมวลผลภาพจากกล่องแต่ละตัว ให้ถูกต้องแม่นยำ โดยระบบติดตามใบหน้าและการเคลื่อนที่ของบุคคลด้วยกล่องวิทัศน์ 2 ตัวที่เป็นต้นแบบนั้น อาศัยวิธีการเทียบมาตรฐานทางสี (Color Calibration) ให้กับกล่องแต่ละตัว ก่อนที่จะนำข้อมูลหลังจากการเทียบมาตรฐานสีนั้นไปประมวลผลเพื่อติดตามใบหน้าต่อไป

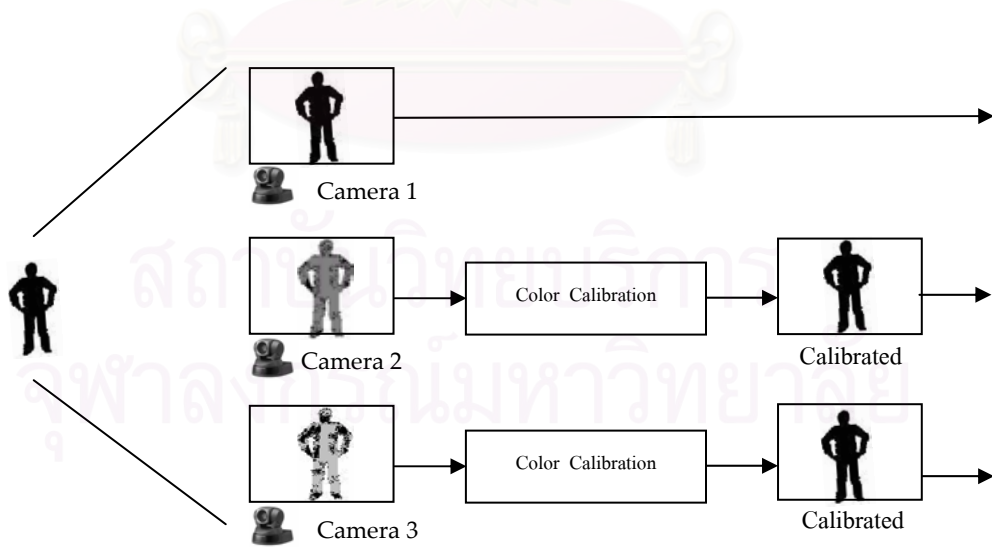
2.5.2 ตัวอย่างสถานการณ์การใช้งานระบบการติดตามใบหน้าและการเคลื่อนที่ของบุคคลโดยใช้กล่อง วิทัศน์มากกว่า 2 ตัว

จากหลักการทั่วไปของการติดตามบุคคลอัตโนมัติระหว่างกล่องหลายตัวในระบบนั้น จะอาศัยการติดตามบุคคลที่ผ่านเข้ามาในระบบกล่องวิทัศน์แต่ละตัวอย่างต่อเนื่อง นั่นคือ จะให้กล่องตัวหนึ่งส่งข้อมูลบุคคลที่มีการเคลื่อนที่ให้กับกล่องตัวอื่นๆ เพื่อทำหน้าที่ติดตามบุคคลอย่างต่อเนื่องโดยอาจจะมีการติดตั้งกล่องให้มีการเชื่อมกัน หรือมีการติดตั้งในตำแหน่งต่างกันโดยไม่มี การเชื่อมกัน ขึ้นอยู่กับความต้องการในการทำงานกัน



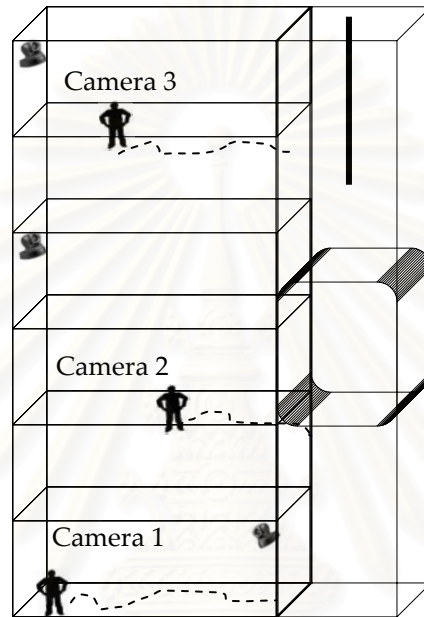
รูปที่ 2.5.1. แสดงตัวอย่างสถานการณ์การติดตั้งระบบการติดตามใบหน้าและการเคลื่อนไหวที่  
ของบุคคล

ด้วยกล้องวิดีโอ 3 ตัว ภายในอาคาร (บริเวณห้องต่างๆ)



รูปที่ 2.5.2 แสดงผลของการเทียบมาตรฐานกล้องวิดีโอ 3 ตัวในระบบ

จากรูปที่ 2.5.1. แสดงตัวอย่างสถานการณ์การติดตั้งระบบการติดตามใบหน้าและการเคลื่อนที่ของบุคคลด้วย กล้องวิดีโอ 3 ตัว ภายในอาคาร ซึ่งกำหนดสมมุติฐานให้ในแต่ละบริเวณที่ติดกล้องนั้น มีความแตกต่างกันทางด้านความสว่าง เช่นเดียวกับสถานการณ์ใน รูปที่ 2.5.3 แสดงตัวอย่างสถานการณ์การติดตั้งระบบการติดตามใบหน้าและการเคลื่อนที่ของบุคคล ด้วยกล้องวิดีโอ 3 ตัว ภายในอาคารบริเวณชั้นต่างๆ ซึ่งทั้งสองสถานการณ์มีจุดประสงค์เพื่อ การติดตามบุคคล



รูปที่ 2.5.3 แสดงตัวอย่างสถานการณ์การติดตั้งระบบการติดตามใบหน้าและการเคลื่อนที่ของบุคคล

ด้วยกล้องวิดีโอ 3 ตัว ภายในอาคาร(บริเวณหน้าลิฟท์)

### 3. ผลสัมฤทธิ์ของโครงการวิจัย

เพื่อให้ทราบถึงระบบการทำงานร่วมกันของกล้องวิดีโอในรูปแบบต่างๆ และทราบถึงตัวอย่างระบบการติดตามใบหน้าและการเคลื่อนที่ของบุคคลด้วยกล้องวิดีโอประเภท PTZ 2 ตัว เพื่อที่จะนำแนวความคิดและปัญหาที่ได้จากการทดลอง เป็นแนวความคิดเบื้องต้นในการ

พัฒนาระบบการติดตามใบหน้าและการเคลื่อนไหวของบุคคลด้วยกล้องวีดีทัศน์ 2 ตัว เพื่อแก้ปัญหา  
ดังกล่าว

ปัญหาอันเนื่องมาจากภาวะทางแสงที่แตกต่างกันในการติดตั้งกล้องแต่ละตัว สามารถ  
แก้ไขได้ด้วยวิธีการเทียบมาตรฐานทางสี (Color Calibration) โดยให้กล้องแต่ละตัว มีมาตรฐาน  
ทางสีเดียวกัน ความผิดพลาดที่เกิดจากการประมวลผลภาพและจากแต่ละกล้องจะลดลง



สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย



## ภาคผนวก ค

### 1. จุดประสงค์ของงานวิจัย

1.1 ศึกษาคุณลักษณะเฉพาะของความสัมพันธ์ระหว่าง อัตราบิต (Bit rate) กับ ค่าความเพี้ยน (Distortion) ของการเข้ารหัส โดยอาศัยความรู้พื้นฐานจากคุณสมบัติทางสถิติของสัญญาณวิดีโอต้นฉบับ (Video source) เพื่อใช้หาแบบจำลองทางคณิตศาสตร์ระหว่างอัตราบิตและค่าความเพี้ยน (Rate-Distortion model : R-D Model) ที่มีความเหมาะสมมากที่สุด เพื่อให้การเข้ารหัสสัญญาณวิดีโอต้นฉบับตามมาตรฐาน H.264/AVC มีประสิทธิภาพสูงสุด

1.2 ศึกษาและพัฒนาทฤษฎีสำหรับกรณีการควบคุมอัตรา โดยอาศัยพื้นฐานของแบบจำลอง R-D ที่ได้ศึกษาตามข้อ 1.1 ในการประมาณค่าหาอัตราบิตและความเพี้ยนของการเข้ารหัส สำหรับการเข้ารหัสในระดับหน่วยพื้นฐาน (Basic Unit) ในมาตรฐานการเข้ารหัสสัญญาณวิดีโอ H.264/AVC

1.3 ศึกษาและพัฒนาทฤษฎีสำหรับกรณีการควบคุมอัตรา ภายใต้เงื่อนไขการประวิงเวลาดำ (Low delay)

### 2. งานวิจัยที่ได้ทำไปแล้ว

2.1 ศึกษาลักษณะการแจกแจงความน่าจะเป็นของการแปลงสัมประสิทธิ์ DCT บล็อก 8x8 เฉพาะค่าสัมประสิทธิ์ AC เปรียบเทียบระหว่างลักษณะการแจกแจงความน่าจะเป็นแบบโคชี (Cauchy Distribution) กับแบบลาปลาเซียน (Laplacian Distribution) แล้วพบว่าสัมประสิทธิ์ AC มีลักษณะการแจกแจงความน่าจะเป็นเป็นแบบโคชี มากกว่าแบบลาปลาเซียน

2.2 วิเคราะห์หาสมการทางคณิตศาสตร์ จาก Lagrange Multiplier Technique โดยอาศัยความสัมพันธ์ระหว่าง Cauchy Rate-Model และ Cauchy Distortion-Model ในการหาสมการที่เหมาะสมที่สุดสำหรับการคำนวณหาค่า Quantization step size เพื่อให้ได้จำนวนบิตตามอัตราบิตเป้าหมายที่กำหนด โดยที่ค่าความเพี้ยนเกิดขึ้นน้อยที่สุด

2.3 นำ Rate-Distortion Model ที่หาได้จากข้อ 2.2 มาประยุกต์ใช้ในกรณีการควบคุมอัตราตามมาตรฐาน H.264/AVC ในระดับหน่วยพื้นฐานย่อย (Basic unit) ภายใต้เงื่อนไขอัตราบิตต่ำ และ

เงื่อนไขการประวิงเวลาต่ำ ซึ่งจากผลการทดลองพบว่าการใช้ Cauchy Rate-Distortion Model สามารถปรับปรุงค่า PSNR ของสัญญาณวิดีโอที่ส่งได้สูงขึ้น, ความแปรปรวนของค่า PSNR น้อยลง และอัตราการกระโดดข้ามเฟรมลดลงเมื่อเทียบกับมาตรฐาน H.264 ภายใต้เงื่อนไขการประวิงเวลาต่ำ

### 3. รายละเอียดของงานวิจัย

งานวิจัยให้ความสำคัญในการแก้ปัญหาหลักของการรับ-ส่งวิดีโอที่ส่งสัญญาณไร้สายอันได้แก่ ความไม่คงที่ของแบนด์วิดท์, การสูญหายของกลุ่มข้อมูล และความเพี้ยนของข้อมูลในระหว่างการส่ง โดยมุ่งเน้นการพัฒนาไปที่กรรมวิธีการควบคุมอัตราสำหรับการเข้ารหัสสัญญาณวิดีโอที่ส่งตามมาตรฐาน H.264/AVC ซึ่งจะทำการหาแบบจำลองทางคณิตศาสตร์ที่เหมาะสม โดยใช้ Lagrange Multiplier Technique เพื่อใช้แสดงความสัมพันธ์ระหว่างอัตราบิต และค่าความเพี้ยนที่เกิดจากการเข้ารหัส แล้วทำการปรับปรุงกลไกในส่วนของกรรมวิธีการควบคุมอัตราใหม่ให้มีความสอดคล้องกับ Rate-Distortion Model , การพิจารณาในส่วนของ Low Delay โดยนำเอาข้อมูลจากสถานะบัฟเฟอร์ของตัวเข้ารหัสมาใช้ในการพิจารณาตัดสินใจ จำนวนบิตเป้าหมายในกรรมวิธีการควบคุมอัตรา ร่วมกับการพิจารณาความซับซ้อน (Complexity) ของแต่ละหน่วยย่อยพื้นฐานซึ่งรายละเอียดของงานวิจัยสามารถสรุปได้ดังนี้

#### ปรับปรุงการเข้ารหัสสัญญาณวิดีโอที่ส่งในส่วนของการควบคุมอัตรา ในมาตรฐาน H.264/AVC โดยใช้ Cauchy Rate-Distortion Model

1. วิเคราะห์สมการทางคณิตศาสตร์เพื่อหาความสัมพันธ์ระหว่างแบบจำลองอัตราบิต และอัตราความเพี้ยนที่เหมาะสมโดยใช้ Lagrange Multiplier Technique

จากบทความเรื่อง “Frame bit allocation for the H.264/AVC video coder via Cauchy-Density –Based rate and Distortion models” สามารถคำนวณหา Rate-Model และ Distortion Model ได้จากสมการ (1) และ (2) ตามลำดับ

$$R = H(Q) = aQ^{-\alpha} \quad (1)$$

$$D(Q) = bQ^\beta \quad (2)$$

โดยที่  $a, \alpha, b, \beta$  คือค่าพารามิเตอร์ของแบบจำลอง

กำหนดให้  $C =$  จำนวนพิกเซลใน basic unit

$H =$  Average rate (bits/pixel) สำหรับการเข้ารหัสเวกเตอร์การเคลื่อนที่ และ coder's header bits.

ดังนั้นจาก (1) ;

$$R_i = Ca_i Q_i^{-\alpha_i} + H_i \quad (3)$$

โดยที่ โดยที่  $i = 1, 2, \dots, N$  และ  $N =$  จำนวน basic unit ภายในเฟรม

จาก(2) , หา Distortion ของ basic unit ที่  $i$  จาก

$$D = \frac{1}{N} \sum_{i=1}^N b_i Q_i^{\beta_i} \quad (4)$$

ทำการหา Optimization ของค่า Quantization สำหรับ Cauchy rate model :

$$Q_1^*, Q_2^*, \dots, Q_N^*, \lambda^* = \arg \min_{Q_1, \dots, Q_N} \frac{1}{N} \sum_{i=1}^N b_i Q_i^{\beta_i}$$

จาก Lagrange Multiplier Technique

$$Q_1^*, Q_2^*, \dots, Q_N^*, \lambda^* = \arg \min_{Q_1, \dots, Q_N} \frac{1}{N} \sum_{i=1}^N b_i Q_i^{\beta_i}$$

$$\text{subject to } \sum_{l=1}^N R_l = R_{MAX} \quad (5)$$

$$Q_1^*, Q_2^*, \dots, Q_N^*, \lambda^* = \arg \min_{Q_1, \dots, Q_N} \frac{1}{N} \sum_{l=1}^N b_l Q_l^{\beta_l} + \lambda \left[ \sum_{l=1}^N R_l - R_{MAX} \right] \quad (6)$$

โดยที่  $R = \text{Constraints bit rate}$  และสมมติให้  $(H_l = H)$  ดังนั้นสมการที่(6) จะสามารถเขียนได้ใหม่ในรูปของ

$$J(Q_1^*, Q_2^*, \dots, Q_N^*, \lambda^*) = \frac{1}{N} \sum_{l=1}^N b_l Q_l^{\beta_l} + \lambda \sum_{l=1}^N (C a_l Q_l^{-\alpha_l}) + \lambda (NH - R_{MAX}) \quad (7)$$

หาค่า  $Q_N^*$  และ  $\lambda^*$  จากการทำ Partial Derivative สมการ (7)

**Case 1 :** หา  $\frac{\partial J}{\partial Q_l^*}$

$$\frac{\partial J}{\partial Q_1^*} = \frac{\beta_1}{N} b_1 (Q_1^*)^{(\beta_1-1)} - (\lambda C a_1 \alpha_1) (Q_1^*)^{-(\alpha_1+1)}$$

$$\frac{\partial J}{\partial Q_2^*} = \frac{\beta_2}{N} b_2 (Q_2^*)^{(\beta_2-1)} - (\lambda C a_2 \alpha_2) (Q_2^*)^{-(\alpha_2+1)}$$

$$\frac{\partial J}{\partial Q_N^*} = \frac{\beta_N}{N} b_N (Q_N^*)^{(\beta_N-1)} - (\lambda C a_N \alpha_N) (Q_N^*)^{-(\alpha_N+1)}$$

$$\therefore \frac{\partial J}{\partial Q_l^*} = \frac{1}{N} (\beta_l b_l Q_l^{(\beta_l-1)}) - \lambda (C a_l \alpha_l Q_l^{-(\alpha_l+1)}) \quad (8)$$

กำหนดให้  $\frac{\partial J}{\partial Q_l^*} = 0$  ดังนั้นจากสมการที่(8),

$$\frac{1}{N}(\beta_l b_l Q_l^{(\beta_l - 1)}) = \lambda (C a_l \alpha_l Q_l^{-(\alpha_l + 1)})$$

$$\frac{Q_l^{(\beta_l - 1)}}{Q_l^{-(\alpha_l + 1)}} = (\lambda CN) \left( \frac{a_l \alpha_l}{\beta_l b_l} \right)$$

$$\therefore Q_l^{(\alpha_l + \beta_l)} = (\lambda CN) \left( \frac{a_l \alpha_l}{\beta_l b_l} \right) \quad (9)$$

พิจารณา (9) , Take ln ทั้งสองข้างของสมการ

$$\ln Q_l^{(\alpha_l + \beta_l)} = \ln (\lambda CN) \left( \frac{a_l \alpha_l}{\beta_l b_l} \right)$$

$$\ln Q_l = \frac{1}{(\alpha_l + \beta_l)} \ln (\lambda CN) \left( \frac{a_l \alpha_l}{\beta_l b_l} \right)$$

$$\therefore Q_l = \left\{ (\lambda CN) \left( \frac{a_l \alpha_l}{\beta_l b_l} \right) \right\}^{\frac{1}{(\alpha_l + \beta_l)}} \quad (10)$$

กำหนดให้ ,  $\frac{1}{(\alpha_l + \beta_l)} = s_l$  และ  $\left( \frac{a_l \alpha_l}{\beta_l b_l} \right) = \varepsilon_l$

ดังนั้น,  $Q_l = ((\lambda CN)(\varepsilon_l))^{s_l} \quad (11)$

**Case 2 :** หา  $\frac{\partial J}{\partial \lambda^*} = 0$

นั่นคือ  $\frac{\partial J}{\partial \lambda} = \sum_{l=1}^N (C a_l Q_l^{-\alpha_l}) + (NH - R_{MAX}) = 0$

ดังนั้น,  $\sum_{l=1}^N (C a_l Q_l^{-\alpha_l}) = (R_{MAX} - NH) \quad (12)$

แทน eq.(11) in eq. (12),

$$\sum_{l=1}^N (Ca_l(\lambda CN \varepsilon_l)^{-s_l \alpha_l}) = (R_{MAX} - NH) \quad (13)$$

ในกรณีนี้กำหนดให้,  $G$  เป็นค่าคงที่  $= \frac{\alpha_l}{\alpha_l + \beta_l}$

$$\sum_{l=1}^N (Ca_l(\lambda CN \varepsilon_l)^{-G}) = (R_{MAX} - NH)$$

$$C^{(1-k)} (\lambda N)^{-k} \sum_{l=1}^N (a_l(\varepsilon_l)^{-G}) = (R_{MAX} - NH)$$

$$(\lambda)^{-k} = N^k C^{(G-1)} \frac{(R_{MAX} - NH)}{\sum_{l=1}^N (a_l(\varepsilon_l)^{-G})}$$

ดังนั้น,

$$\lambda = \left( \frac{\sum_{l=1}^N (a_l(\varepsilon_l)^{-G})}{N^G C^{(G-1)} (R_{MAX} - NH)} \right)^{\frac{1}{G}} \quad (14)$$

แทนสมการที่(14) ลงในสมการ (11), ดังนั้นจะสามารถคำนวณหา  
ค่าควอนไทล์เซชันที่เหมาะสมที่สุดได้จากสมการที่ (15)

$$Q_l = \left( \left( \frac{\sum_{l=1}^N (a_l(\varepsilon_l)^{-G})}{N^G C^{(G-1)} (R_{MAX} - NH)} \right)^{\frac{1}{G}} CN(\varepsilon_l) \right)^{s_l}$$

$$\therefore Q_l = \left( \varepsilon_l \left( \sum_{l=1}^N (a_l(\varepsilon_l)^{-G}) \frac{C}{R_{MAX} - NH} \right)^{\frac{1}{G}} \right)^{s_l} \quad (15)$$

โดยที่  $a_l$ ,  $\alpha_l$  และ  $b_l$ ,  $\beta_l$  คือค่าพารามิเตอร์ของแบบจำลองของแต่ละหน่วยย่อยพื้นฐานที่  $l$ th ตามลำดับ

2. ในการหาค่าพารามิเตอร์ของแบบจำลองอัตราบิดและความเพี้ยนแบบโคชี ( $a_{j-1}$ ,  $\alpha_{j-1}$  และ  $b_{j-1}$ ,  $\beta_{j-1}$ ) โดยใช้ข้อมูลทางสถิติของอัตราบิด และความเพี้ยนของหน่วยพื้นฐานย่อยที่ถูกเข้ารหัสไปแล้ว ซึ่งเมื่อพิจารณาสมการที่ (3) และ (4) สามารถเขียนให้อยู่ในรูปแบบฟังก์ชันเชิงเส้นได้ดังสมการที่ (1) และ (2)

$$\ln(R_i(n_{i,j})) = \ln(a_{j-1}) - \alpha_{j-1} \ln(Q_i(n_{i,j})) \quad (16)$$

$$\ln(MSE_i(n_{i,j})) = \ln(b_{j-1}) + \beta_{j-1} \ln(Q_i(n_{i,j})) \quad (17)$$

และจากเทคนิคการวิเคราะห์การถดถอยแบบเชิงเส้น จะสามารถคำนวณหาค่าพารามิเตอร์ของแบบจำลองในหน่วยพื้นฐานย่อยที่  $l$  ได้ด้วยสมการที่ (18)-(21)

$$\ln(a_{j-1}) = \frac{\left( \sum_k^{l-1} \ln(R_k) \right) \left( \sum_k^{l-1} (\ln(Q_k))^2 \right) - \left( \sum_k^{l-1} \ln(Q_k) \ln(R_k) \right)}{(l-k) \left( \sum_k^{l-1} (\ln(Q_k))^2 \right) \left( \sum_k^{l-1} \ln(Q_k) \right)^2} \quad (18)$$

$$\alpha_{j-1} = \frac{\left( \sum_k^{l-1} \ln(Q_k) \right) \left( \sum_k^{l-1} (\ln(R_k)) \right) - (l-k) \left( \sum_k^{l-1} \ln(Q_k) \ln(R_k) \right)}{(l-k) \left( \sum_k^{l-1} (\ln(Q_k))^2 \right) \left( \sum_k^{l-1} \ln(Q_k) \right)^2} \quad (19)$$

$$\ln(b_{j-1}) = \frac{\left( \sum_k^{l-1} \ln(MSE_k) \right) \left( \sum_k^{l-1} (\ln(Q_k))^2 \right) - \left( \sum_k^{l-1} \ln(Q_k) \ln(MSE_k) \right)}{(l-k) \left( \sum_k^{l-1} (\ln(Q_k))^2 \right) \left( \sum_k^{l-1} \ln(Q_k) \right)^2} \quad (20)$$

$$\beta_{j-1} = \frac{(l-k) \left( \sum_k^{l-1} \ln(Q_k) \ln(MSE_k) \right) - \left( \sum_k^{l-1} \ln(Q_k) \right) \left( \sum_k^{l-1} \ln(MSE_k) \right)}{(l-k) \left( \sum_k^{l-1} (\ln(Q_k))^2 \right) \left( \sum_k^{l-1} \ln(Q_k) \right)^2} \quad (21)$$

โดยที่  $R_k$ ,  $Q_k$  และ  $MSE_k$  หมายถึง จำนวนบิตที่ใช้เข้ารหัส, ระดับขั้นของการควอนไทซ์ และ ค่าความเพี้ยนของสัญญาณในหน่วยพื้นฐานย่อยที่ถูกเข้ารหัสไปแล้ว ตั้งแต่หน่วยย่อยพื้นฐานที่  $k$  ถึง  $l-1$

3. ทำการ Implement ในส่วนของ Rate control ลงในมาตรฐานการเข้ารหัสสัญญาณวิดีโอ ทัศน์ H.264 โดยใช้ Cauchy Rate-Distortion model แทนการใช้ Quadratic rate model ซึ่งเป็นแบบจำลองอัตราบิตที่ได้จากการวิเคราะห์ PDF ของ AC-Coefficients ของลำดับภาพเป็นแบบ Laplacian distribution โดยแบ่งขั้นตอนการทำงานออกเป็น GOP –Layer, Frame- Layer และ Basic Unit – Layer ดังนี้

- การควบคุมอัตราในระดับ GOP Layer: คำนวณหาจำนวนบิตรวมของเฟรมที่เหลืออยู่ใน GOP ( $B_i(j)$ ) เมื่อเฟรมที่  $j$  ของ GOP ที่  $i$  ถูกเข้ารหัสตั้งสมการที่ (22)-(23)

$$T_{GOP} = \frac{u(n_{i,j})}{F_r} \times N_{GOP} \quad (22)$$

$$T_r(n_{i,j}) = T_{GOP} - \sum_{j=0}^{j-1} b(n_{i,j}) \quad (23)$$

- ควบคุมอัตราในระดับ Frame Layer: คำนวณหาจำนวนบิตเป้าหมายของแต่ละเฟรม

คำนวณหาค่าควอนไทซ์เริ่มต้นสำหรับ  $I$  เฟรม, และ  $P$  เฟรมแรก ( $Q_I^0, Q_P^0$ ) โดยพิจารณาจาก Channel bandwidth และ ขนาดของ GOP (ตามมาตรฐาน H.264) และคำนวณหาอัตราบิตเป้าหมายของแต่ละ  $P$  เฟรมดังนี้

1) คำนวณหาระดับบัพเฟอร์เป้าหมาย ของแต่ละ  $P$  เฟรมจากสมการ



$$B_c(n_{i,j}) = B_c(n_{i,j-1}) + (A(n_{i,j-1}) - (\frac{u(n_{i,j-1})}{F_r})) \quad (24)$$

2) คำนวณหาจำนวนบิตเป้าหมายของแต่ละ P เฟรม โดยพิจารณาจาก 2 กรณีคือ

- พิจารณาจากจำนวนบิตที่เหลืออยู่ของแต่ละเฟรม

$$\hat{f}(n_{i,j}) = \frac{T_r(n_{i,j})}{N_{GOP} - j} \quad (25)$$

- พิจารณาจากระดับบัพเฟอร์เป้าหมาย ซึ่งจำนวนบิตเป้าหมายสามารถคำนวณได้จากสมการ

$$f(n_{i,j}) = \hat{f}(n_{i,j}) + \omega V_S - B_c(n_{i,j}) \quad (26)$$

โดยที่  $\omega$  เป็นค่าถ่วงน้ำหนัก โดยกำหนดให้มีค่าเท่ากับ 0.8

- **ควบคุมอัตราในระดับ Basicunit Layer** : แบ่งออกเป็น 2 ขั้นตอนคือ

1) **Pre-encoding stage** : คำนวณหาค่าระดับขั้นของการควอนไทซ์สำหรับแต่ละหน่วยย่อยพื้นฐาน

**กรณี 1** สำหรับหน่วยย่อยพื้นฐานแรกของแต่ละ P เฟรม ค่าควอนไทซ์ที่ไต่เซชันพารามิเตอร์จะพิจารณาจากค่าควอนไทซ์ที่ไต่เซชันพารามิเตอร์ของทุกๆ หน่วยย่อยพื้นฐานใน P เฟรมก่อนหน้า

$$QP_{1,i}(j) = \text{Avg}QP_i(j-1) \quad (27)$$

**กรณี 2** เมื่อจำนวนบิตที่เหลืออยู่มีค่าน้อยกว่าศูนย์ ค่าควอนไทซ์ที่ไต่เซชันพารามิเตอร์จำคำนวณจาก

$$QP_{1,i}(j) = QP_{1-1,i}(j) + 2 \quad (28)$$

**กรณี 3** คำนวณหาค่าระดับชั้นของการควอนไทซ์จาก Cauchy Rate – Distortion Model ในสมการที่ (15) โดยพิจารณาจาก complexity factor ของแต่ละหน่วยย่อยพื้นฐาน ตามขั้นตอนต่อไปนี้

1) คำนวณค่า complexity factor ของแต่ละหน่วยย่อยพื้นฐานจาก

$$\gamma(j,l) = \frac{Res_{var}(j,l)}{Ave\_Res_{var}(j-1)} \quad (29)$$

แล้วคำนวณหาจำนวนบิตเป้าหมายที่ใช้ในการเข้ารหัสแต่ละ P เฟรมใหม่ โดยพิจารณาจากค่า complexity factor ดังสมการต่อไปนี้

$$f_v(n_{i,j}) = \begin{cases} f(n_{i,j}) & \text{if } \gamma(j,l) \leq 1 \\ 1.1 \times f(n_{i,j}) & \text{if } \gamma(j,l) > 1 \end{cases} \quad (30)$$

2) คำนวณหาค่าระดับชั้นของการควอนไทซ์ของแต่ละหน่วยย่อยพื้นฐานโดยใช้ค่าโมเดลพารา-มิเตอร์  $a_{j-1}$ ,  $\alpha_{j-1}$  และ  $b_{j-1}$ ,  $\beta_{j-1}$ ,  $R_{MAX}(j)$  ของแต่ละหน่วยย่อยพื้นฐาน

## 2) Post Encoding Stage

- Update ค่าพารามิเตอร์โมเดล  $a_{j-1}$ ,  $\alpha_{j-1}$  และ  $b_{j-1}$ ,  $\beta_{j-1}$  ด้วยสมการที่ (18)-(21)

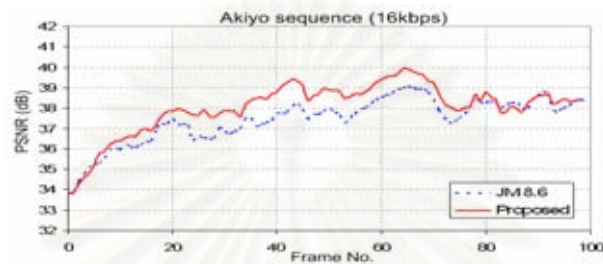
## ผลการทดลองกรณีไม่คิด Low delay

ทำการทดลองใช้ Cauchy Rate Distortion Model ที่วิเคราะห์ได้ ตามรายละเอียดที่ผ่านมานมาตรฐาน H.264/AVC (JM 8.6) โดยมีพารามิเตอร์ของการเข้ารหัสสัญญาณวิดีโอ ดังนี้

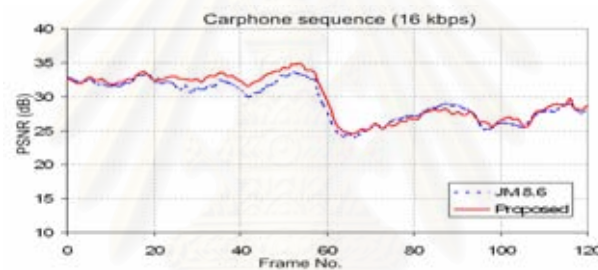
- 1) ทำการเข้ารหัสแบบ IPPP...
- 2) เข้ารหัสที่อัตราบิตต่างๆ คือ 16 kbps , 32 kbps, 64 kbps, 128 kbps และ 256 kbps
- 3) ใช้ลำดับภาพทดสอบจำนวน 7 ลำดับภาพ (Akiyo, Carphone, Claire, Foreman, News, Silent และ Missam\_Suzie)
- 4) สัญญาณวิดีโอที่ใช้เป็นแบบ QCIF ขนาด 144 x 176 พิกเซล

5) ในการทดลองใช้มาตรฐาน H.264 เวอร์ชัน JM 8.6 ในระดับ Main profile

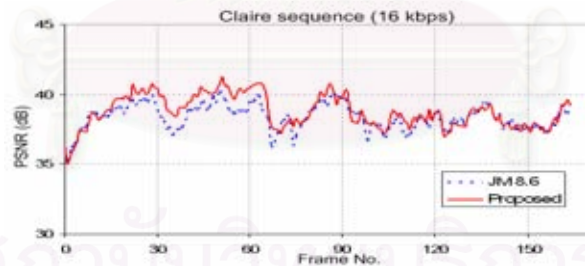
ซึ่งจากผลการทดลองพบว่า Cauchy Rate Distortion Model ที่นำเสนอ นั้นสามารถปรับปรุงค่า PSNR ได้สูงขึ้นเมื่อเทียบกับการเข้ารหัสตามมาตรฐาน H.264 โดยที่ีมีความแปรปรวนของค่า PSNR และ processing times ต่ำกว่า JM 8.6 ซึ่งสามารถแสดงรายละเอียดของผลการทดลองได้ในตารางที่ 1. และกราฟรูปที่ 1-8



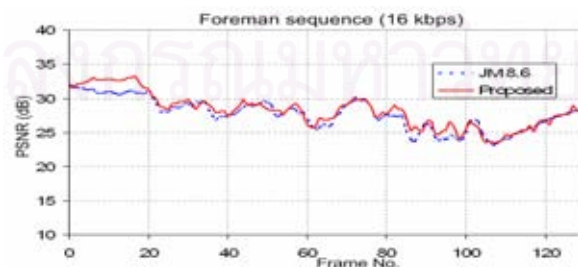
(a)



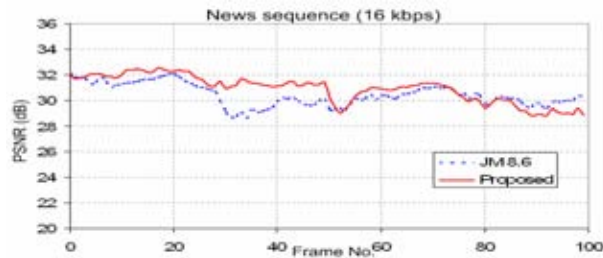
(b)



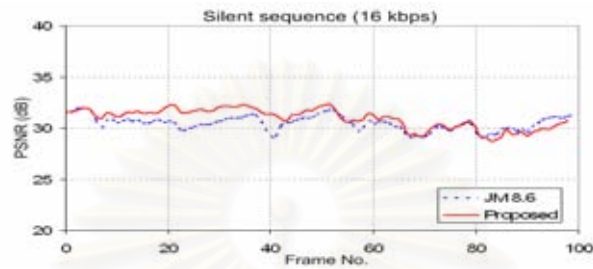
(c)



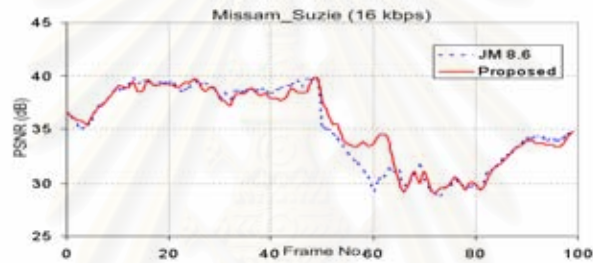
(d)



(e)

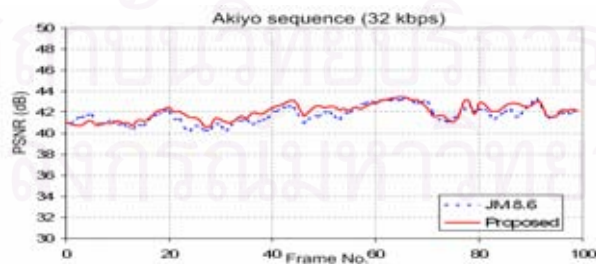


(f)

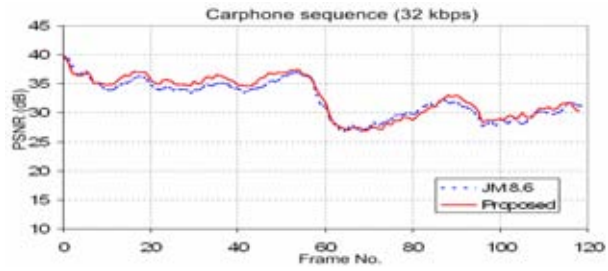


(g)

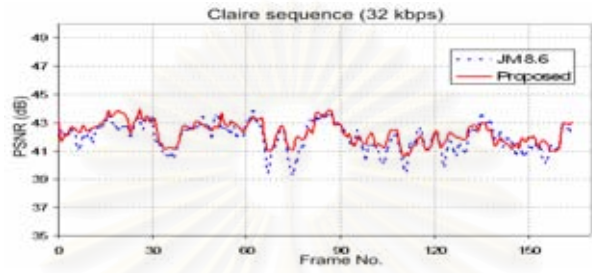
รูปที่ 1. แสดงค่า PSNR ของแต่ละเฟรมเปรียบเทียบกันระหว่างกรรมวิธีการควบคุมอัตราตามมาตรฐาน H.264 เวอร์ชัน JM8.6 กับ กรรมวิธีที่นำเสนอ ของแต่ละลำดับภาพทดสอบที่อัตราบิต 16 kbps



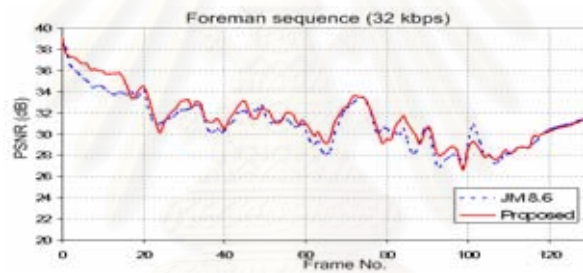
(a)



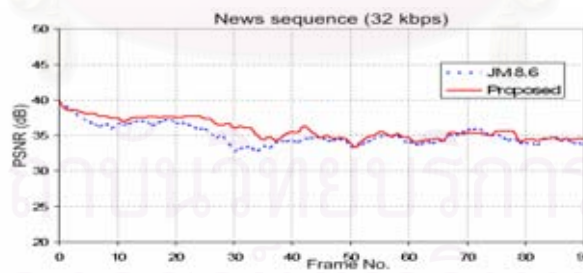
(b)



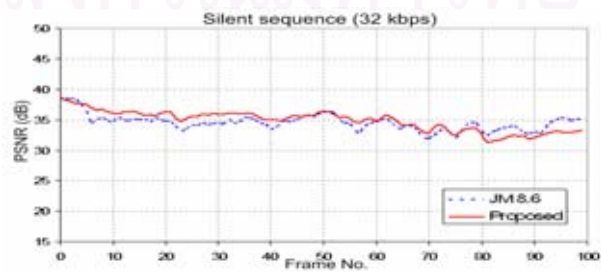
(c)



(d)



(e)



(f)

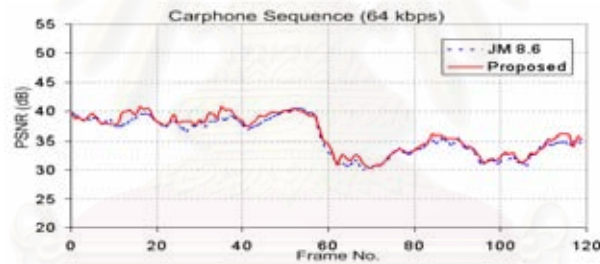


(g)

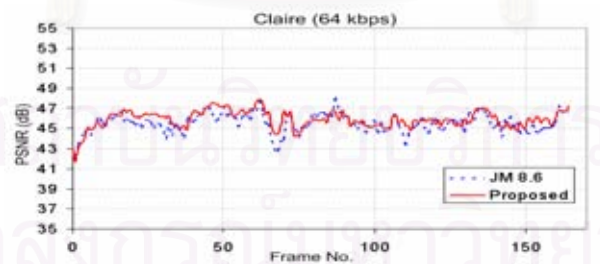
รูปที่ 2. แสดงค่า PSNR ของแต่ละเฟรมเปรียบเทียบกันระหว่างกรรมวิธีการควบคุมอัตราตามมาตรฐาน H.264 เวอร์ชัน JM8.6 กับ กรรมวิธีที่นำเสนอ ของแต่ละลำดับภาพทดสอบที่อัตราบิต 32 kbps



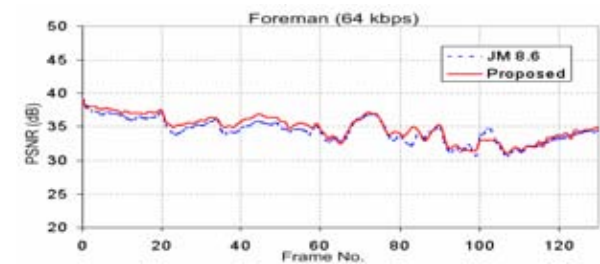
(a)



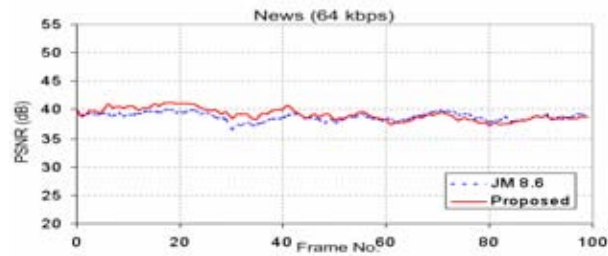
(b)



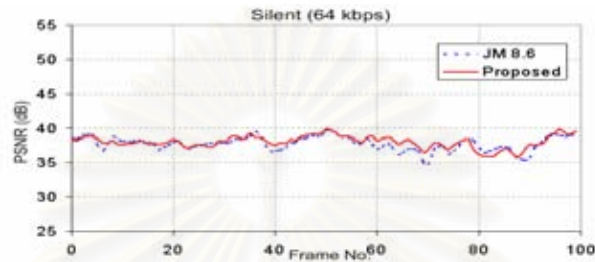
(c)



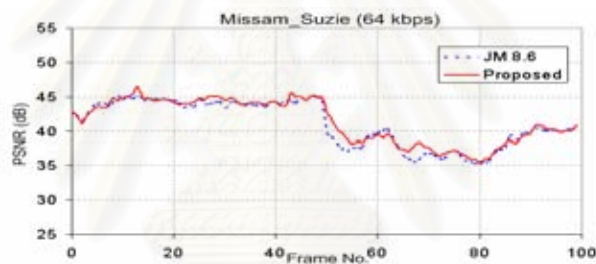
(d)



(e)

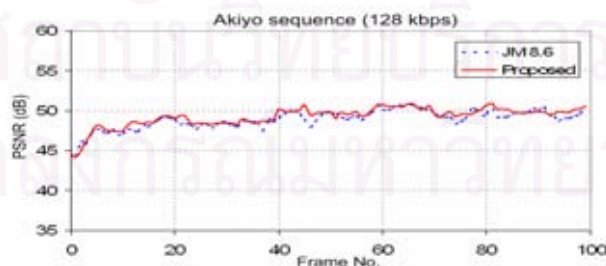


(f)

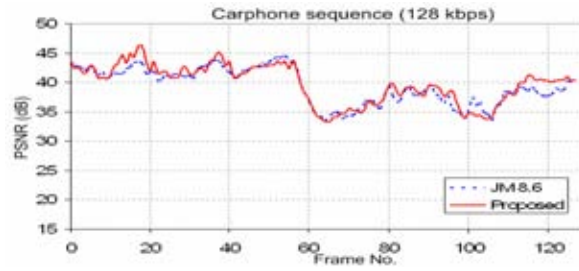


(g)

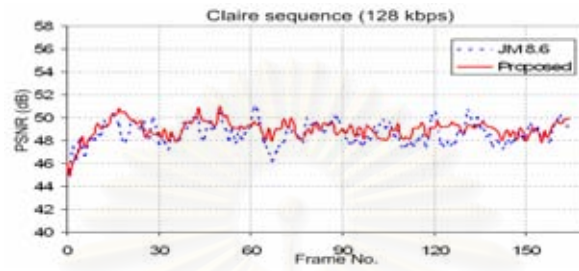
รูปที่ 3. แสดงค่า PSNR ของแต่ละเฟรมเปรียบเทียบกันระหว่างกรรมวิธีการควบคุมอัตราตามมาตรฐาน H.264 เวอร์ชัน JM8.6 กับ กรรมวิธีที่นำเสนอ ของแต่ละลำดับภาพทดสอบที่อัตราบิต 64 kbps



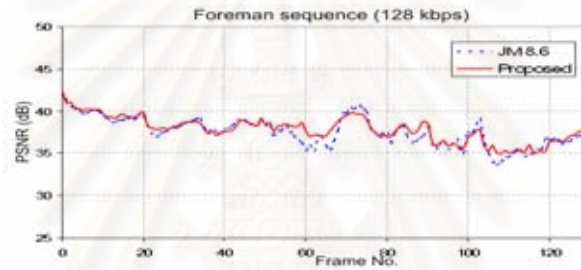
(a)



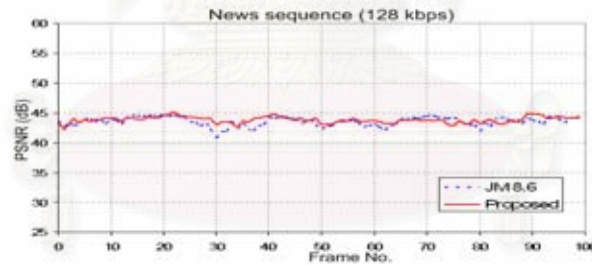
(b)



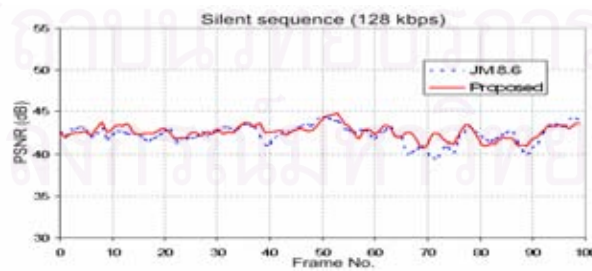
(c)



(d)



(e)



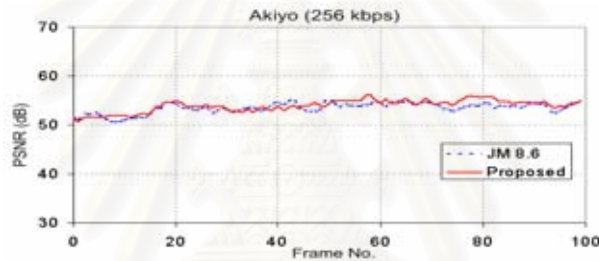
(f)



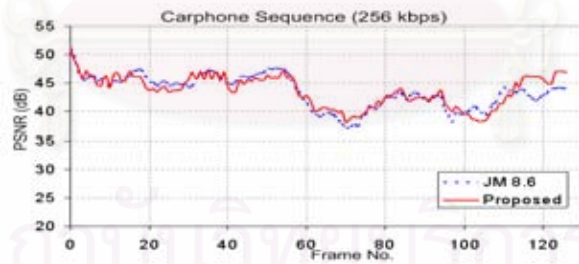


(g)

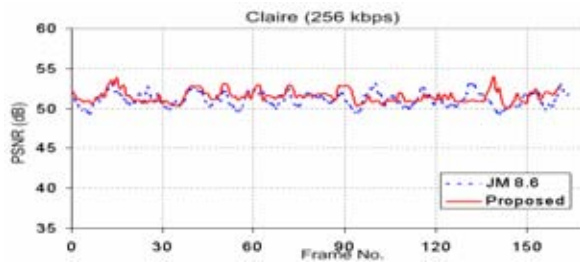
รูปที่ 4. แสดงค่า PSNR ของแต่ละเฟรมเปรียบเทียบกันระหว่างกรรมวิธีการควบคุมอัตราตามมาตรฐาน H.264 เวอร์ชัน JM8.6 กับ กรรมวิธีที่นำเสนอ ของแต่ละลำดับภาพทดสอบที่อัตราบิต 128 kbps



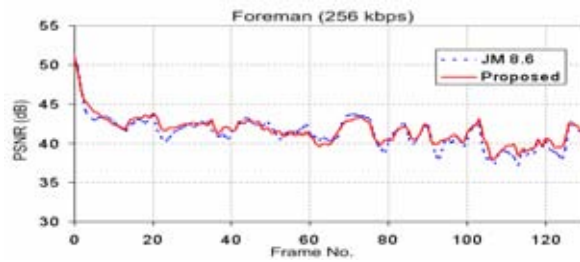
(a)



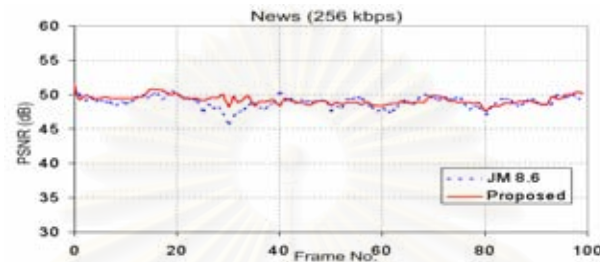
(b)



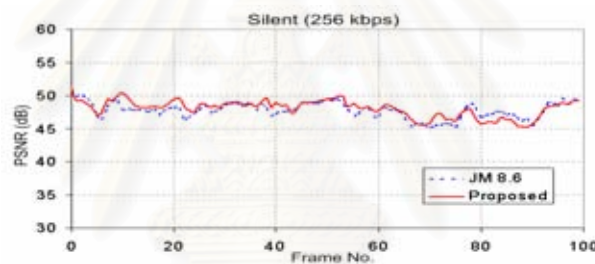
(c)



(d)



(e)

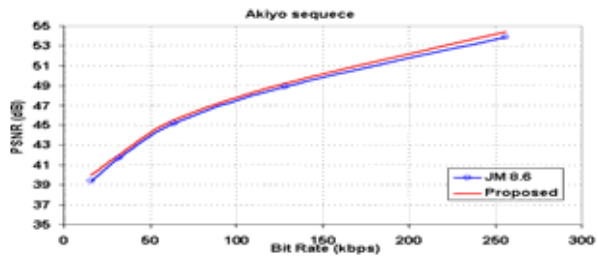


(f)

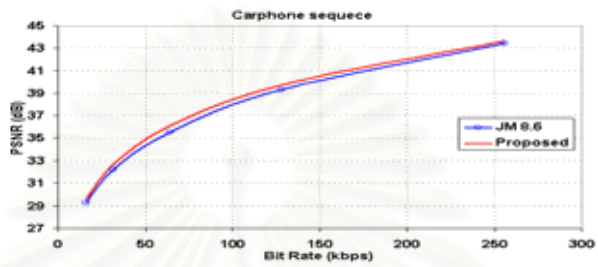


(g)

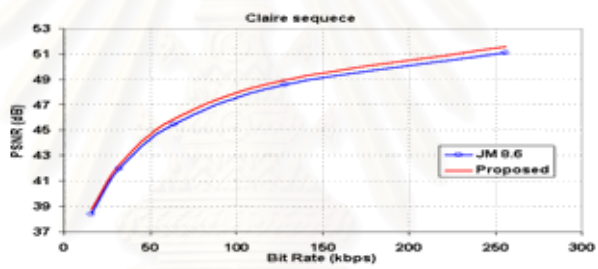
รูปที่ 5. แสดงค่า PSNR ของแต่ละเฟรมเปรียบเทียบกันระหว่างกรรมวิธีการควบคุมอัตรา ตามมาตรฐาน H.264 เวอร์ชัน JM8.6 กับ กรรมวิธีที่นำเสนอ ของแต่ละลำดับภาพทดสอบที่อัตรา บิต 256 kbps



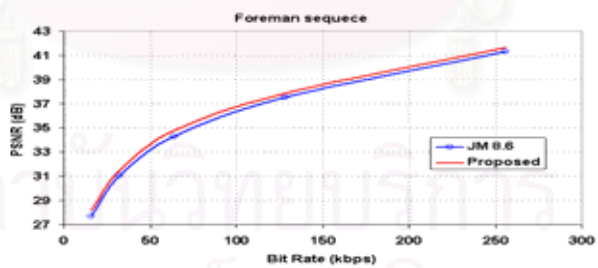
(a)



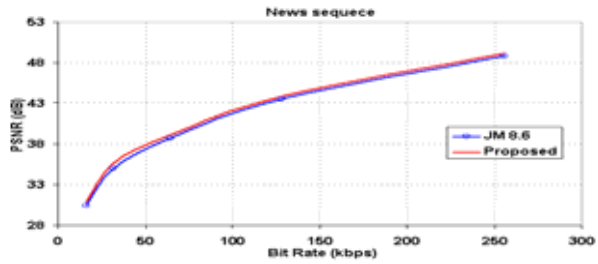
(b)



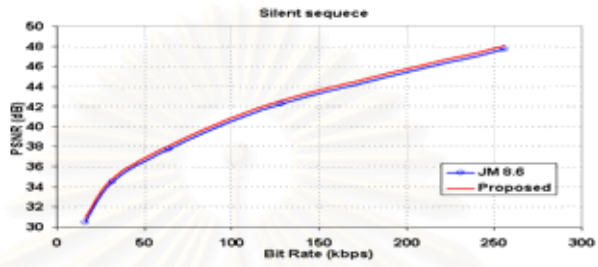
(c)



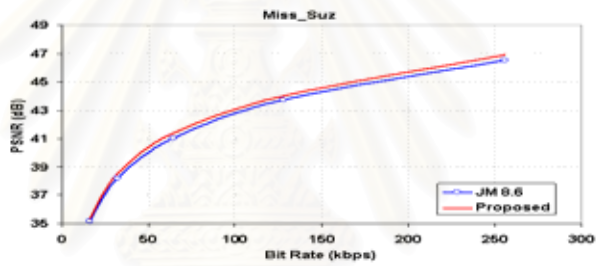
(d)



(e)



(f)



(g)

รูปที่ 6. แสดงค่า PSNR เฉลี่ยของแต่ละลำดับภาพทดสอบที่แต่ละอัตราบิตทดสอบ (16-256 kbps)

สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย



(a1)



(a2)



(b1)



(b2)

รูปที่ 7. ภาพถอดรหัสทดสอบ Missam\_Suzie ที่ผ่านการเข้ารหัสด้วยกรรมวิธีควบคุมอัตรา  
ตามาตรฐาน JM8.6

เปรียบเทียบกับภาพที่ได้จากกรรมวิธีควบคุมอัตราที่นำเสนอ

(a1) เฟรมที่ 62 ของกรรมวิธีควบคุมอัตราที่นำเสนอ (PSNR= 34.56 dB) ที่อัตราบิต 16  
kbps

(a2) เฟรมที่ 62 ของกรรมวิธีการควบคุมอัตราตามมาตรฐาน JM 8.6 (PSNR= 30.55 dB)  
ที่อัตราบิต 16 kbps

(b1) เฟรมที่ 71 ของกรรมวิธีควบคุมอัตราที่นำเสนอ (PSNR= 32.90 dB) ที่อัตราบิต 32  
kbps

(b2) เฟรมที่ 71 ของกรรมวิธีการควบคุมอัตราตามมาตรฐาน JM 8.6 (PSNR= 28.86 dB)  
ที่อัตราบิต 32 kbps



(a1)



(a2)



(b1)



(b2)

รูปที่ 8. ภาพถอดรหัสทดสอบ News ที่ผ่านการเข้ารหัสด้วยกรรมวิธีควบคุมอัตราตามาตรฐาน JM8.6

เปรียบเทียบกับภาพที่ได้จากกรรมวิธีควบคุมอัตราที่นำเสนอ.

(a1) เฟรมที่ 50 ของกรรมวิธีควบคุมอัตราที่นำเสนอ (PSNR= 38.67 dB) ที่อัตราบิต 16 kbps

(a2) เฟรมที่ 50 ของกรรมวิธีควบคุมอัตราตามมาตรฐาน JM 8.6 (PSNR= 37.67 dB) ที่อัตราบิต 16 kbps

(b1) เฟรมที่ 63 ของกรรมวิธีควบคุมอัตราที่นำเสนอ (PSNR= 39.91 dB) ที่อัตราบิต 16 kbps

(b2) เฟรมที่ 63 ของกรรมวิธีควบคุมอัตราตามมาตรฐาน JM 8.6 (PSNR= 38.92 dB) ที่อัตราบิต 16 kbps

ตารางที่ 1: เปรียบเทียบประสิทธิภาพของกรรมวิธีการควบคุมตามมาตรฐาน H.264 กับกรรมวิธี  
ควบคุมอัตราที่นำเสนอ

Bit rate	Sequence	Average PSNR (dB)		PSNR Std.		Processing Times (ms)		Bit Rate (kbps)		PPSNR Gain (dB)
		JM 8.6	Proposed	JM 8.6	Proposed	JM8.6	Proposed	JM8.6	Proposed	
kbps 16	Akiyo	39.40	40.03	1.12	1.26	233.25	224.83	16.08	16.10	+0.63
	Claire	38.36	38.76	0.95	0.92	222.15	223.88	16.08	16.07	+0.40
	Silent	30.45	30.93	0.71	0.75	223.59	224.02	16.07	16.18	+0.48
	News	30.41	30.90	0.87	0.90	219.43	212.81	16.07	16.03	+0.49
	Carphone	29.27	29.62	2.73	2.42	232.34	235.41	16.02	16.13	+0.35
	Foreman	27.68	28.21	2.24	2.06	229.94	230.07	16.11	16.07	+0.53
	Miss_Suz	35.17	35.37	3.67	3.40	272.06	230.34	15.99	16.15	+0.20
kbps 32	Akiyo	41.71	42.02	0.81	0.77	233.88	212.77	31.98	32.06	+0.31
	Claire	41.96	42.27	0.98	0.80	215.52	215.50	31.92	31.97	+0.31
	Silent	34.50	34.85	1.30	1.25	213.88	212.19	32.08	32.05	+0.35
	News	34.97	35.66	1.39	1.37	212.75	213.90	32.39	32.12	+0.69
	Carphone	32.24	32.69	3.10	2.98	219.49	220.15	32.32	32.11	+0.45
	Foreman	31.04	31.48	2.29	2.18	225.80	226.10	32.07	32.03	+0.44
	Miss_Suz	38.23	38.54	3.81	3.56	255.56	248.26	31.95	32.23	+0.31
kbps 64	Akiyo	45.22	45.50	1.33	1.37	231.79	211.99	63.87	64.11	+0.28
	Claire	45.48	45.85	0.95	0.89	219.53	220.08	63.85	64.05	+0.37
	Silent	37.74	38.05	0.98	0.87	219.01	218.95	63.98	64.10	+0.31
	News	38.68	39.13	0.73	0.72	215.22	215.43	64.31	64.23	+0.45
	Carphone	35.56	36.10	3.01	3.04	223.25	223.81	64.98	64.12	+0.54
	Foreman	34.32	34.80	1.71	1.62	227.60	227.69	63.98	64.08	+0.48

	Miss_Suz	41.02	41.39	3.28	3.16	232.85	230.14	63.63	64.19	+0.37
kbps 128	Akiyo	48.89	49.24	1.22	1.21	231.24	212.45	127.91	127.86	+0.35
	Claire	48.60	48.96	0.99	0.81	221.19	221.30	127.32	127.88	+0.36
	Silent	42.26	42.55	1.05	0.88	219.68	220.06	127.58	128.03	+0.29
	News	43.52	43.83	0.77	0.50	218.94	217.16	128.14	128.19	+0.31
	Carphone	39.29	39.73	3.10	3.28	221.20	220.48	127.96	128.06	+0.44
	Foreman	37.55	37.84	1.69	1.50	227.96	227.52	127.98	128.03	+0.29
	Miss_Suz	43.70	44.01	2.72	2.48	234.97	232.62	127.59	128.05	++0.31
kbps 256	Akiyo	53.37	53.87	1.09	1.12	239.29	207.69	256.16	256.08	+0.50
	Claire	51.09	51.54	0.90	0.77	221.76	221.54	254.93	255.27	+0.45
	Silent	47.73	48.05	1.27	1.23	212.14	212.15	255.37	256.04	+0.32
	News	48.80	49.16	0.87	0.64	219.55	218.34	256.71	256.33	+0.36
	Carphone	43.44	43.64	2.91	2.75	226.18	226.47	255.93	256.06	+0.22
	Foreman	41.33	41.64	1.96	1.73	226.34	226.73	254.28	255.37	+0.31
	Miss_Suz	46.53	46.89	2.16	1.93	236.72	235.41	253.25	256.69	+0.36

4. ทำการ Implement ในส่วนของ Rate control ภายใต้เงื่อนไขการประวิงเวลาต่ำ ลงในมาตรฐานการเข้ารหัสสัญญาณวีดิทัศน์ H.264 โดยใช้ Cauchy Rate-Distortion model แทนการใช้ Quadratic rate model ซึ่งเป็นแบบจำลองอัตราบิตที่ได้จากการวิเคราะห์ PDF ของ AC-Coefficients ของลำดับภาพเป็นแบบ Laplacian distribution โดยแบ่งขั้นตอนการทำงานออกเป็น GOP –Layer , Frame- Layer และ Basic Unit – Layer ดังนี้

- การควบคุมอัตราในระดับ GOP Layer: คำนวณหาจำนวนบิตรวมของเฟรมที่เหลืออยู่ใน GOP ( $B_i(j)$ ) เมื่อเฟรมที่  $j$  ของ GOP ที่  $i$  ถูกเข้ารหัสดังสมการที่ (22)-(23)

- ควบคุมอัตราในระดับ Frame Layer: คำนวณหาจำนวนบิตเป้าหมายของแต่ละเฟรม



คำนวณหาค่าควอนท์ไตรซ์เริ่มต้นสำหรับ I เฟรม, และ P เฟรมแรก ( $Q_I^0, Q_P^0$ ) โดยพิจารณาจาก Channel bandwidth และ ขนาดของ GOP (ตามมาตรฐาน H.264) และคำนวณหาอัตราบิตเป้าหมายของแต่ละ P เฟรมดังนี้

1) คำนวณหาระดับบัพเฟอ์เป้าหมาย ของแต่ละ P เฟรมจากสมการที่ (24)

2) คำนวณหาจำนวนบิตเป้าหมายของแต่ละ P เฟรม โดยพิจารณาจาก 2 กรณีคือ

- พิจารณาจากจำนวนบิตที่เหลืออยู่ของแต่ละเฟรมตามสมการที่ (25)

- พิจารณาจากระดับบัพเฟอ์เป้าหมาย ซึ่งจำนวนบิตเป้าหมายสามารถคำนวณได้จากสมการที่ (26)

- **ควบคุมอัตราในระดับ Basicunit Layer** : แบ่งออกเป็น 2 ขั้นตอนคือ

1) Pre-encoding stage : คำนวณหาระดับขั้นของการควอนท์ไตรซ์สำหรับแต่ละหน่วยย่อยพื้นฐาน

**กรณี 1** สำหรับหน่วยย่อยพื้นฐานแรกของแต่ละ P เฟรม ค่าควอนท์ไตรซ์เซชันพารามิเตอร์จะพิจารณาจากค่าควอนท์ไตรซ์เซชันพารามิเตอร์ของทุกๆ หน่วยย่อยพื้นฐานใน P เฟรมก่อนหน้าดังในสมการที่ (27)

**กรณี 2** เมื่อจำนวนบิตที่เหลืออยู่มีค่าน้อยกว่าศูนย์ ค่าควอนท์ไตรซ์เซชันพารามิเตอร์จำคำนวณจากสมการที่ (28)

**กรณี 3** คำนวณหาระดับขั้นของการควอนท์ไตรซ์จาก Cauchy Rate – Distortion Model ในสมการที่ (15) โดยพิจารณาจาก complexity factor ของแต่ละหน่วยย่อยพื้นฐาน ตามขั้นตอนต่อไป

1) คำนวณค่า complexity factor ของแต่ละหน่วยย่อยพื้นฐานจากสมการที่ (29) แล้วคำนวณหาจำนวนบิตเป้าหมายที่ใช้ในการเข้ารหัสแต่ละ P เฟรมใหม่โดยพิจารณาจากค่า complexity factor ดังสมการต่อไปนี้

$$f_v(n_{i,j}) = \gamma(j,l) \times f(n_{i,j}) \quad (31)$$

- 2) เพื่อลดปัญหาการกระโดดข้ามเฟรม ในงานวิจัยนี้จึงได้ทำการปรับจำนวนบิตเป้าหมายของแต่ละ P เฟรมใหม่ โดยพิจารณาจากระดับบัพเฟอร์ เช่นในกรณีที่ระดับบัพเฟอร์มีค่าสูงเกินกว่า 80 % ของขนาดบัพเฟอร์สูงสุด จำนวนบิตเป้าหมายจะถูกปรับลดลงมา 10 % แต่ในกรณีที่ระดับบัพเฟอร์มีค่าต่ำกว่าศูนย์ (หมายถึง buffer underflow) จำนวนบิตเป้าหมายจะถูกปรับให้มีค่าสูงขึ้น 10 % ดังสมการต่อไปนี้

$$R_{MAX}(j) = \eta \times f_v(n_{i,j}) \quad (32)$$

$$\eta = \begin{cases} 1.10 & \text{if } B_c(n_{i,j}) \leq 0 \\ 0.90 & \text{if } B_c(n_{i,j}) \geq 0.8 \times V_S \\ 1.00 & \text{otherwise} \end{cases} \quad (33)$$

- 3) เพื่อลด PSNR fluctuation และป้องกันการเกิด Buffer overflow จำนวนบิตเป้าหมายของแต่ละ P เฟรม จะถูกจำกัดที่ค่าสูงสุด ต่ำสุด ดังสมการที่ (34)-(35)

$$Lower\_bound\_R_{MAX} = MAX \left\{ R_{MAX}(j), \left( \phi \times \frac{u(n_{i,j})}{F_r} \right) \right\} \quad (34)$$

$$Upper\_bound\_R_{max} = MIN \left\{ R_{MAX}(j), \left( \gamma \times \frac{u(n_{i,j})}{F_r} \right) \right\} \quad (35)$$

- 4) คำนวณหาค่าระดับชั้นของการควอนไทซ์ของแต่ละหน่วยย่อยพื้นฐานโดยใช้ค่าโมเดลพารา-มิเตอร์  $\alpha_{j-1}$ ,  $\alpha_{j-1}$  และ  $\beta_{j-1}$ ,  $\beta_{j-1}$ ,  $R_{MAX}(j)$  ของแต่ละหน่วยย่อยพื้นฐาน โดยที่ค่าควอนไทซ์เซชันพารามิเตอร์จะถูกปรับเปลี่ยนไปตามสมการต่อไปนี้

$$\text{if } B_S(n_{i,j}) \geq 1.2 \times B_S$$

$$QP_{i,j}(j) = MAX(QP_{initial} + 5, QP_{i,j}(j) + 3)$$

$$\text{else if } B_S(n_{i,j}) \leq 0.2 \quad (36)$$

$$QP_{l,i}(j) = \text{MAX}(QP_{\text{initial}} - 2, QP_{l,i}(j))$$

else

$$QP_{l,i}(j) = \text{MAX}(QP_{\text{initial}} - 1, QP_{l,i}(j))$$

โดยที่  $B_S(n_{i,j})$  สามารถคำนวณได้จากสมการที่ (37)

$$B_S(n_{i,j}) = B_c(n_{i,j}) + \sum_{k=1}^l b(n_{i,j-k}) \quad (37)$$

## 2) Post Encoding Stage

- Update ค่าพารามิเตอร์โมเดล  $a_{j-1}$ ,  $\alpha_{j-1}$  และ  $b_{j-1}$ ,  $\beta_{j-1}$  ด้วยสมการที่ (18)-

(21)

### ผลการทดลองกรณีคิด Low delay

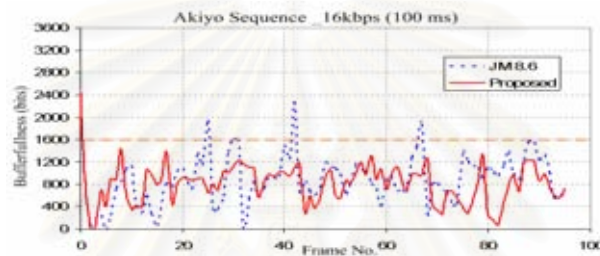
ทำการทดลองใช้ Cauchy Rate Distortion Model ที่วิเคราะห์ได้ ตามรายละเอียดที่ผ่าน มาบนมาตรฐาน H.264/AVC (JM 8.6) โดยมีพารามิเตอร์ของการเข้ารหัสสัญญาณวิดีโอที่สนใจดังนี้

- 1) ทำการเข้ารหัสแบบ IPPP...
- 2) เข้ารหัสที่อัตราบิตต่างๆ คือ 16 kbps, 32 kbps, 64 kbps, 128 kbps และ 256 kbps
- 3) ใช้ลำดับภาพทดสอบจำนวน 7 ลำดับภาพ (Akiyo, Carphone, Claire, Foreman, News, Silent และ Missam\_Suzie)
- 4) สัญญาณวิดีโอที่ใช้เป็นแบบ QCIF ขนาด 144 x 176 พิกเซล

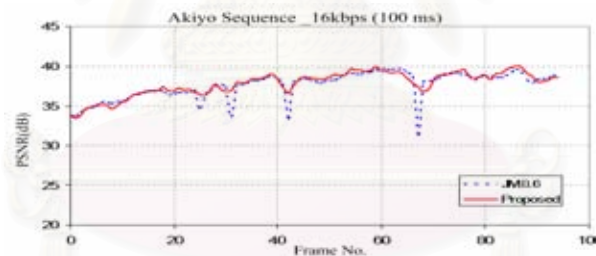
5) ทำการทดลองโดยปรับให้มีขนาดบัฟเฟอร์ของตัวเข้ารหัสที่แตกต่างกันไป จำนวน 4 ค่าด้วยกันเพื่อให้ได้ค่า Delay times ที่แตกต่างกันคือ 100 ms, 150 ms, 200 ms และ 400 ms ตามลำดับ

6) ในการทดลองใช้มาตรฐาน H.264 เวอร์ชัน JM 8.6 ในระดับ Main profile

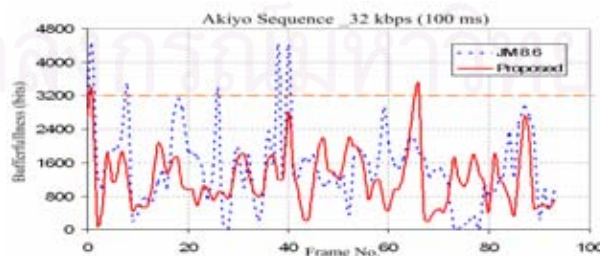
ซึ่งจากผลการทดลองพบว่า Cauchy Rate Distortion Model ที่นำเสนอภายใต้เงื่อนไขการประวิงเวลาดำเนินสามารถปรับปรุงค่า PSNR ได้สูงขึ้น และมีอัตราการกระโดดข้ามเฟรมลดลง เมื่อเทียบกับการเข้ารหัสตามมาตรฐาน H.264 โดยที่มีความแปรปรวนของค่า PSNR และ processing times ต่ำกว่า JM 8.6 ซึ่งสามารถแสดงรายละเอียดของผลการทดลองได้ในตารางที่ 2. และกราฟรูปที่ 9 ดังนี้



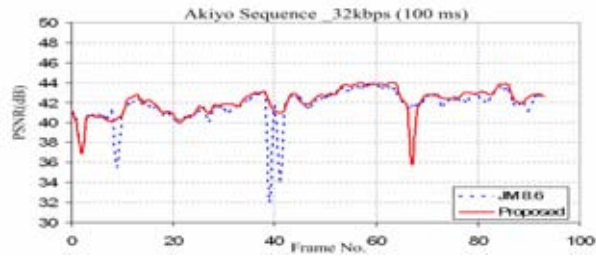
(a1)



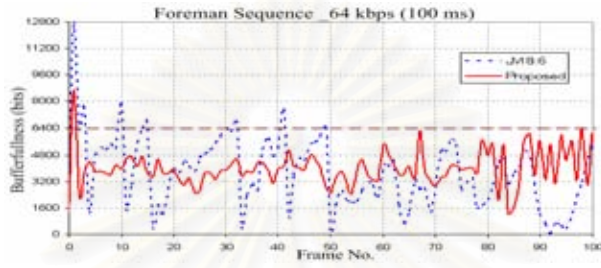
(b1)



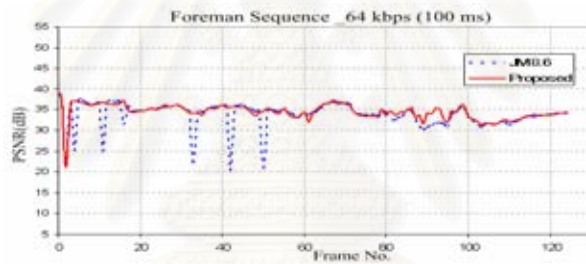
(a2)



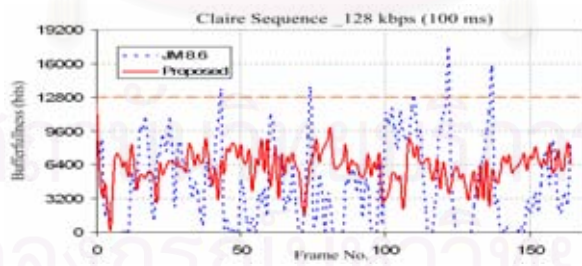
(b2)



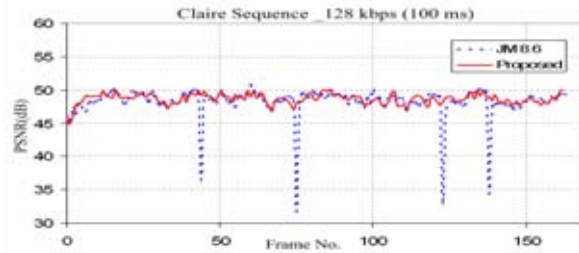
(a3)



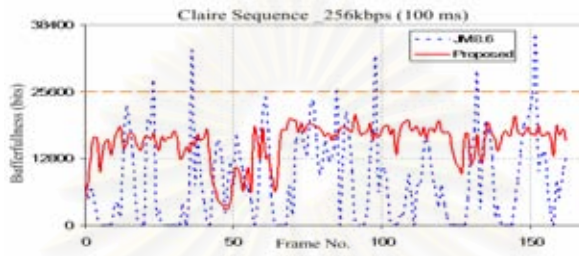
(b3)



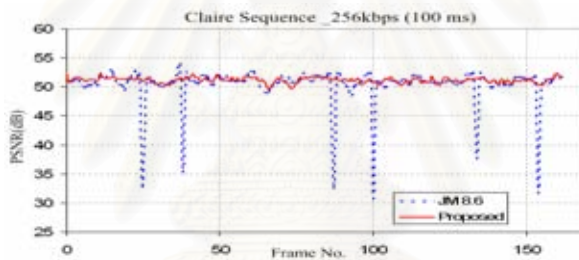
(a4)



(b4)



(a5)



(b5)

รูปที่ 9. ผลการทดลองโดยใช้ลำดับภาพทดสอบจำนวน 6 ลำดับภาพทดสอบเปรียบเทียบกันระหว่างกรรมวิธีควบคุมอัตราที่นำเสนอกับกรรมวิธีการควบคุมอัตราตามมาตรฐาน H.264 ที่อัตราบิต 16-256 kbps ภายใต้เงื่อนไขการประวิงเวลาต่ำที่ 100 ms.

(a) ระดับการใช้บัพเฟอร์ของแต่ละเฟรม โดยมีขนาดบัพเฟอร์สูงสุดแสดงด้วยเส้นประ

(b) ค่าเฉลี่ย PSNR ของแต่ละเฟรม

ตารางที่ 2: เปรียบเทียบประสิทธิภาพของกรรมวิธีการควบคุมตามมาตรฐาน H.264 กับกรรมวิธี  
ควบคุมอัตราที่นำเสนอ ภายใต้เงื่อนไขการประวิงเวลาต่ำ 100 ms

Bit rate	Sequence	Average PSNR (dB)		PSNR Std.		No. frame skipped		Bit Rate (kbps)		Processing Times (ms)		PSNR Gain(dB)	of frame skipped reduction
		JM 8.6	Proposed	JM 8.6	Proposed	JM 8.6	Proposed	JM 8.6	Proposed	JM 8.6	Proposed		
16 kbps	Akiyo	37.59	37.80	1.69	1.53	4	0	16.04	16.04	271.02	227.37	+0.21	100%
	Claire	38.02	38.37	2.24	1.65	16	4	15.98	16.04	218.89	216.43	+0.35	75%
	Silent	32.47	32.83	1.82	1.25	12	5	16.06	16.05	239.15	219.51	+0.36	58%
	News	30.52	30.86	1.22	0.98	5	4	15.97	16.06	288.47	222.68	+0.34	20%
	Carphone	29.01	29.70	3.07	2.71	13	5	16.02	16.01	219.74	220.45	+0.69	62%
	Foreman	26.73	27.40	3.45	2.55	17	9	15.53	16.03	252.00	240.30	+0.67	47%
32 kbps	Akiyo	41.73	42.08	1.81	1.37	5	2	32.00	32.09	218.92	222.70	+0.35	60%
	Claire	41.86	42.32	2.01	1.36	4	1	31.90	32.07	219.35	215.04	+0.46	75%
	Silent	38.60	39.96	2.95	1.66	11	2	31.32	32.04	222.93	220.48	+1.36	82%
	News	39.25	39.70	2.41	2.01	5	3	31.87	32.28	213.38	214.75	+0.45	40%
	Carphone	31.87	32.51	3.98	3.65	10	6	31.91	32.25	220.50	209.87	+0.64	40%
	Foreman	30.14	31.04	3.88	3.39	11	6	30.70	31.04	226.85	227.46	+0.90	45%
64 kbps	Akiyo	45.23	45.52	1.36	1.44	0	0	63.95	63.52	228.78	222.31	+0.29	-
	Claire	45.37	45.81	1.58	1.18	1	0	63.46	63.69	219.37	218.89	+0.44	100%
	Silent	36.96	37.39	3.34	2.47	9	4	64.09	64.14	223.89	221.53	+0.43	56%
	News	38.77	39.24	1.22	0.83	1	0	63.70	63.73	249.58	230.12	+0.47	100%
	Carphone	36.28	36.88	3.79	3.13	7	3	63.40	64.16	221.85	223.10	+0.60	57%
	Foreman	33.84	34.55	3.09	1.89	7	1	63.95	64.09	231.13	225.69	+0.71	86%
128 kbps	Akiyo	48.98	49.27	1.25	1.26	0	0	127.66	127.35	225.53	212.12	+0.29	-
	Claire	48.24	48.94	2.47	0.79	4	0	124.24	127.84	220.16	222.63	+0.70	100%

	Silent	43.89	44.57	2.48	0.82	2	0	127.92	128.70	213.40	215.02	+1.68	100%
	News	44.94	45.50	1.73	0.68	1	0	127.40	128.01	237.98	236.49	+0.69	100%
	Carphone	38.94	39.31	3.82	3.16	3	1	126.53	128.90	212.21	210.45	+0.37	67%
	Foreman	37.50	37.88	1.95	1.66	1	0	127.57	127.66	222.47	209.83	+0.38	100%
256 kbps	Akiyo	52.92	53.72	2.28	1.09	2	0	253.45	256.26	213.00	214.14	+0.80	100%
	Claire	50.41	51.13	3.49	0.47	6	0	245.20	253.59	222.60	220.77	+0.72	100%
	Silent	47.58	48.06	2.62	1.11	1	0	256.33	256.45	221.62	215.63	+0.48	100%
	News	48.67	49.29	2.13	2.17	1	1	255.86	255.15	236.46	230.41	+0.62	-
	Carphone	42.82	43.61	4.78	3.47	5	2	253.5	257.99	222.10	219.54	+0.79	60%
	Foreman	41.07	41.86	3.33	2.50	3	1	254.92	255.91	212.83	218.67	+0.79	67%

#### 4. ผลสัมฤทธิ์ของโครงการวิจัย

พัฒนาระบบซอฟต์แวร์สำหรับจำลองการรับ-ส่ง สัญญาณวีดิทัศน์ตามมาตรฐาน H.264 /AVC ในส่วนของกรรมวิธีการควบคุมอัตรา โดยใช้ Cauchy Rate Distortion Model และทำการปรับปรุงกรรมวิธีการควบคุมอัตราใหม่ในมาตรฐาน H.264 ให้สอดคล้องกับ Cauchy Rate Distortion Model ที่วิเคราะห์ได้ โดยพิจารณาร่วมกับการนำไปประยุกต์ใช้กับการส่งสัญญาณวีดิทัศน์แบบเวลาจริงที่เป็นแบบ Low delay และ Low bit rate

สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย



## เอกสารอ้างอิง

1. ISO/IEC JTC1/SC29/WG11, Test Model 5, 1993
2. Video Coding for Low Bit Rate Communications, ITU-T, ITU-T Recommendation H.263, ver. 1, 1995.
3. T. Sikora, "The MPEG-4 video standard verification model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 1, pp. 19–31, Feb. 1997.
4. "Draft ITU-T recommendation and final draft international standard of joint video specification (ITU-T Rec.H.264/ISO/IEC 14 496-10 AVC," in JVT of ISO/IEC MPEG and ITU-T VCEG, JVTG050, 2003.
5. T.Berger, "Rate Distortion Theory," Englewood Cliffs, NJ: Prentice-Hall, 1984.
6. C.E.Shannon, "A mathematical theory of communications," *Bell System Tech. Journal*, vol.27, pp.397-423; 623-656, July and Oct. 1948.
7. T.M.Cover and J.A.Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
8. R.M.Gray, *Source Coding Theory*. Norwell, MA: Kluwer, 1990.
9. Stanislaw H. Zak and Edwin K.P. Chong, *An introduction to optimization*. New York: Wiley, 2001.
10. R.E.Bellman, *Dynamic programming*. Princeton, NJ: Princeton University Press, 1957.
11. G.D.Forney, "The Viterbi algorithm," *Proc. IEEE*, vol.61, pp.268-278, Mar.1973.
12. W.K.Pratt, *Digital Image Processing*. New York: Wiley, 1978, ch.10.
13. F. Muller, "Distribution shape of two-dimensional DCT coefficients natural images," *Electron. Lett.*, vol. 29, no. 22, pp. 1935–1936, Oct. 1993.
14. S. R.Smooth and R.A.Lowe, "Study of DCT coefficients distributions," in *Proc. SPIE*, Jan. 1996, pp.403-411.
15. T.Eude, R. Grisel, H. Cherifi, and R. Debrie, "On the distribution of the DCT coefficients," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, vol.5, Apr. 1994, pp. 365-368.
16. L.-J. Lin, A. Ortega, and C.-C. J. Kuo, "Cubic spline approximation of rate and distortion functions for MPEG video," *Vis. Commun. Image Process.*, Mar. 1996.

17. Y.Altunbasak and N.Kamaci, "Frame bit allocation for the H.264/AVC video coder via Cauchy-Density –Based rate and Distortion models," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, No. 8, Aug. 2005.
18. R.J.Freund and W.J.Wilson, "Regression Analysis: Statistical Modeling of a Response Variable," New York: Academic, 1998, pp. 39–41.
19. J.Wei and B.H.Soong, "A new Rate – Distortion Model for Video Transmission Using Multiple Logarithmic Functions," *IEEE Signal Processing Letters*, vol.11, No. 8, Aug, 2004
20. I.H.Shin and Y.L.Lee, " Rate control using linear rate  $\rho$ -model for H.264," *Signal processing: Image Communication* 19(2004) 341-352.
21. N. Eiamjumrus and S. Aramvith, "Cauchy Based Rate-Distortion Optimization Model for H.264 Rate Control," *Proceeding of IEEE Asia Pacific Conference on Circuits and Systems (APCCAS) 2006*, Singapore, December 2006.
22. N. Eiamjumrus and S. Aramvith, "New Rate-Control Scheme based on Cauchy Rate-Distortion Optimization Model for H.264 Video Coding," *Proceeding of IEEE International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS) 2006*, Japan, December 2006.

สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย