

โครงการวิจัยย่อยลำดับที่ 9

เรื่อง การศึกษาและพัฒนากระบวนการเข้ารหัสและถอดรหัสสัญญาณวิดีโอตามมาตรฐาน MPEG4 เพื่อส่ง
บนระบบเครือข่ายสื่อสารไร้สาย ปีที่ 4 : การเข้ารหัสวิดีโอที่สนับสนุนอัตราบิตต่ำเพื่อการส่งไร้สายและงาน
ประยุกต์เชิงวัตถุ
(Low Bit-Rate Video Coding for Wireless Transmission and Object-based Applications)

ผู้รับผิดชอบโครงการย่อย ผศ. ดร. สุภาวดี อร่ามวิทย์

1. ปัญหาที่ทำการวิจัยและความสำคัญของปัญหา

ในระบบการสื่อสารโทรคมนาคมยุคต่อไป มีขีดความต้องการรับและส่งสัญญาณหลากหลายรูปแบบมากขึ้น ในรูปแบบของสื่อประสม ซึ่งรวมถึง ข้อมูลตัวอักษร เสียงพูด ออดิโอ ภาพนิ่ง และ วิดิทัศน์ เพื่อรองรับการใช้งานด้านต่างๆ ที่ขยายตัวอย่าง รวดเร็ว อย่างไรก็ตาม ช่องสัญญาณที่มีการกำหนดให้ระบบสื่อสารโทรคมนาคม แต่ละระบบมีขนาดจำกัด ดังนั้น ปริมาณของสัญญาณจำนวนมากที่เกิดขึ้น จะทำให้เกิดความล่าช้าหรือ การประวิงเวลา รวมทั้ง อาจเกิดการสูญหายของข้อมูล ซึ่งอาจก่อให้เกิดความเสียหายในการสื่อสารได้ ในสื่อ ประสมทั้งหมดที่กล่าวมา สัญญาณวิดีโอ จัดเป็นสัญญาณที่มีปริมาณ ข้อมูลสูงที่สุด โดยเฉพาะอย่างยิ่ง ถ้าต้องการ ได้ภาพที่มีความละเอียดสูง

ดังนั้นงานวิจัยที่เกี่ยวข้องกับการลดรหัสสัญญาณวิดีโอ จึงได้รับความสนใจทั้งในแวดวงการศึกษา และ วงการอุตสาหกรรม รวมถึงมีการตั้งองค์กรขึ้นเพื่อกำหนดมาตรฐานลดรหัสสัญญาณวิดีโอเพื่อ ใช้ในการเก็บลงสื่อประเภทต่างๆ และสำหรับการส่งผ่านช่องสัญญาณประเภทต่างๆ แนวทางการวิจัยจะมุ่งเน้น ในส่วนของ การนำเสนอกรรมวิธีการเข้ารหัสให้มีอัตราการบีบอัดข้อมูลที่สูง โดยยังสามารถรักษาระดับคุณภาพของสัญญาณ และการปรับปรุงประสิทธิภาพในการส่งผ่านช่องสัญญาณ ผ่าน โครงข่ายสื่อสารทั้งแบบมีสายและ ไร้สาย

ปัจจุบัน มาตรฐานการลดรหัสสัญญาณวิดีโอที่สนับสนุนอัตราบิตและการประวิงเวลาต่ำที่มีการใช้ในการประยุกต์ระบบการประชุมสัมมนาทางวิดีโอและการส่งสัญญาณวิดีโอผ่านเครือข่ายสื่อสาร ได้แก่ มาตรฐาน H.263 [10.1] MPEG-4 [10.2] และ มาตรฐาน H.264/MPEG-4 Part 10 [10.3] เป็นต้น มาตรฐาน MPEG-4 เป็นมาตรฐานการบีบอัดสัญญาณวิดีโอเชิงวัตถุ ซึ่งให้ประสิทธิภาพการบีบอัดที่สูง เมื่อเทียบกับมาตรฐานก่อนหน้า มาตรฐาน MPEG-4 ได้ถูกนำมาประยุกต์เชิงพาณิชย์ ไม่ว่าจะเป็น ตัวบีบอัดสัญญาณวิดีโอในโทรศัพท์ไร้สายยุคที่ 3 ระบบการประชุมสัมมนาทางวิดีโอ 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 ทำให้การสร้างและ

ประยุกต์ระบบการสื่อสารสัญญาณวิดีโอผ่านเครือข่ายสื่อสารต้องมีการพึ่งพิงการนำเข้าทรัพยากรจากต่างประเทศเป็นหลัก ทำให้ต้องเสียค่าใช้จ่ายเป็นจำนวนมาก ดังนั้นความสามารถในการใช้และพัฒนาองค์ความรู้ทางด้านการเข้ารหัสสัญญาณวิดีโอ และการสื่อสารผ่านเครือข่ายไร้สายจึงเป็นสิ่งสำคัญในการนำประเทศก้าวไปสู่การพึ่งพาตนเองทางด้านวิทยาศาสตร์และเทคโนโลยี

2. วัตถุประสงค์

- 2.1 เพื่อสร้างองค์ความรู้ทางกระบวนการเข้ารหัสสัญญาณวิดีโออัตราบิตต่ำ กรรมวิธีการเข้ารหัส สัญญาณวิดีโอแยกส่วน การติดตามวัตถุโดยใช้เทคนิคทางคอมพิวเตอร์วิทัศน์ การจัดสรรบิตให้ กับวัตถุที่สนใจ รวมถึง เทคนิคการเพิ่มความทนทานของสัญญาณวิดีโอเมื่อส่งผ่านช่องสัญญาณไร้สาย
- 2.2 เพื่อนำองค์ความรู้ที่ได้มาใช้ในการสร้างจริงงานประยุกต์เชิงวัตถุ เช่น ระบบการประชุมสัมมนาทางวิดีโอ

3. แนวเหตุผล และทฤษฎี

มาตรฐานการเข้ารหัสสัญญาณวิดีโอ MPEG-4 [10.2] และ H.264/MPEG-4 Part 10 [10.3] ได้มีการกำหนดโครงสร้างการเข้ารหัสและถอดรหัส สัญญาณวิดีโอ ซึ่งประกอบด้วย ส่วนการลดรหัสทรานส์ฟอร์ม (transform coding) การควอนไทซ์ (quantization) การลดรหัสเอนโทรปี (entropy coding) และการประมาณและชดเชยการเคลื่อนที่ (motion estimation and motion compensation) ในปัจจุบัน การเข้ารหัสเชิงวัตถุ และการส่งวิดีโอผ่าน เครือข่ายไร้สาย มีความสำคัญสำหรับงานประยุกต์ทางวิดีโอ เช่น ระบบประชุมสัมมนาทางวิดีโอ ระบบการประชุมเสมือน (virtual meeting) การส่งวิดีโอผ่านมือถือ (Video on Mobile) และ ระบบตรวจการณ์ทางไกล เป็นต้น ในส่วนการเข้ารหัสเชิงวัตถุ กรรมวิธีการแยกส่วนภาพ (image segmentation) เป็นส่วนประกอบสำคัญ แต่อย่างไรก็ดี ไม่ได้มีการกำหนดกรรมวิธีไว้ในมาตรฐานการเข้ารหัสสัญญาณวิดีโอ ดังนั้น จึงมีการวิจัยทางนี้อย่างกว้างขวาง องค์ความรู้ที่สำคัญในการแยกส่วนภาพ คือ กรรมวิธีการประมวลผลภาพ ซึ่งประกอบด้วยทฤษฎีทางพื้นฐานทาง การตรวจจับขอบ การตรวจจับการเปลี่ยนแปลงระหว่างภาพ การหาค่าทางสถิติของลักษณะเด่นของภาพ เช่น สี พื้นผิว และรูปร่าง และ กรรมวิธีทาง Morphology เป็นต้น ทั้งนี้ ในโครงการวิจัยที่ผ่านมา ได้นำเสนอกรรมวิธีการเข้ารหัสสัญญาณวิดีโอแยกส่วน โดยใช้การแยกส่วนสัญญาณวิดีโอแบบทนทานโดยใช้วิธีการทางสถิติเชิงพารามตริกแบบปรับตัวได้ โดยปรับปรุงกรรมวิธีการลบฉากหลังให้มีประสิทธิภาพที่ดีขึ้น [10.4-10.7] และมีการนำเสนอกรรมวิธีการตรวจหาใบหน้าและมือมนุษย์ ซึ่งสามารถนำไปประยุกต์ใช้ได้กับงานประยุกต์การส่งวิดีโอภาษามือ [10.8-10.9] ในโครงการวิจัยปีนี้ จะเพิ่มประสิทธิภาพของกรรมวิธีการแยกส่วนภาพที่ได้พัฒนาขึ้น ซึ่งจะต้องมีความซับซ้อนต่ำ เพื่อให้เหมาะสมกับการนำไปประยุกต์ในส่วนตัวเข้ารหัสสัญญาณวิดีโอ เพื่อใช้ในการส่งสัญญาณวิดีโอเวลาจริง รวมถึง พัฒนาเทคนิคการตรวจหาและติดตามใบหน้า โดยใช้กล้องวิดีโอ 2 ตัว ในการส่งสัญญาณวิดีโอเวลาจริงผ่านเครือข่ายไร้สายนั้น

ปัจจัยสำคัญ คือ ความน่าเชื่อถือของระบบการส่งสัญญาณวิทยุคลื่นเวลาจริง ซึ่งส่งผลถึง คุณภาพของ สัญญาณวิทยุที่ปลายทาง ว่ามีความผิดพลาดเป็นที่สังเกตเห็นชัดเจนหรือไม่ ส่วนประกอบสำคัญที่เกี่ยวข้อง คือ กรรมวิธีการจัดสรรบิตของตัวเข้ารหัส และ เทคนิคในการเข้ารหัสเพื่อให้สัญญาณวิทยุที่มีความทนทานต่อความผิดพลาดของ สัญญาณวิทยุในการส่งผ่านโครงข่ายไร้สาย ในโครงการนี้จะได้ ทำการศึกษาในส่วนนี้ เพื่อให้สัญญาณ วิทยุที่มีคุณภาพที่ดี และมีความผิดพลาดน้อยที่สุด ซึ่งถือเป็น ปัจจัยที่สำคัญต่อระบบการสื่อสารสัญญาณวิทยุ

4. ขั้นตอนการวิจัย

โครงการวิจัยช่วงปีที่ 4 นี้ได้ถูกแบ่งออกเป็นการทำงาน 2 ส่วน โดยจะเป็นการสานต่องานวิจัย ทางด้าน การเข้ารหัสสัญญาณวิทยุที่อัตราบิตต่ำ จากโครงการในปีที่ 1 - 3 มีรายละเอียดของวิธีการ ทำงานในแต่ละส่วน และระยะเวลา ดังต่อไปนี้

ส่วนที่ 1 การสานต่อการพัฒนาเทคนิคการเข้ารหัสสัญญาณวิทยุเชิงวัตถุ ที่มีประสิทธิภาพสูง มี ความซับซ้อนน้อย และสามารถลดปริมาณข้อมูล ได้มากขึ้น ในส่วนของการแยกส่วนภาพ และการ ตรวจสอบ และ ติดตามใบหน้า

ส่วนที่ 2 การสานต่อส่วนการพัฒนาเทคนิคในการเข้ารหัสเพื่อให้สัญญาณวิทยุมีความทนทานต่อ ความ ผิดพลาดของสัญญาณวิทยุ ในการส่งผ่านโครงข่ายไร้สาย รวมถึง วิธีแก้ไขและปกปิดความ ผิดพลาด การเข้ารหัสแบบสเกล เพื่อเพิ่มคุณภาพของสัญญาณวิทยุ เพื่อทำการส่งบนโครงข่ายสื่อสาร ไร้สาย

ทั้งนี้ มีแผนการดำเนินงานและระยะเวลาดังต่อไปนี้

การดำเนินงาน	เดือนที่ 1-3	เดือนที่ 4-6	เดือนที่ 7-9	เดือนที่ 10-12
1. ส่วนการพัฒนาเทคนิค การเข้ารหัส สัญญาณวิทยุเชิงวัตถุ	←————→			
2. ส่วนการพัฒนาเทคนิคในการเข้ารหัส เพื่อให้สัญญาณวิทยุมีความทนทานต่อ ความ ผิดพลาดของสัญญาณวิทยุ ใน การส่งผ่านโครงข่ายไร้สาย	←————→			
3. ส่วนการทดสอบประสิทธิภาพของ ขั้นตอนวิธีที่นำเสนอ	←————→			
4. ส่วนการจัดทำรายงาน				←————→

5. ขอบเขต และเป้าหมาย

การวิจัยในโครงการจะครอบคลุมถึง

- 5.1 กระบวนการเข้ารหัสสัญญาณวีดิทัศน์อัตราบิตต่ำ กรรมวิธีการเข้ารหัสสัญญาณวีดิทัศน์แยกส่วน การติดตามวัตถุโดยใช้เทคนิคทางคอมพิวเตอร์วิทัศน์ การจัดสรรบิตให้กับวัตถุที่สนใจ รวมถึงเทคนิคการเพิ่มความทนทานของสัญญาณวีดิทัศน์เมื่อส่งผ่านช่องสัญญาณ ไร้สาย
- 5.2 การนำขั้นตอนวิธีของการเข้ารหัสที่พัฒนาได้ มาประยุกต์ควบคู่กับ วิธีแก้ไขและปกปิดความผิดพลาด เพื่อเพิ่มคุณภาพของสัญญาณวีดิทัศน์ เพื่อทำการส่งบนโครงข่ายสื่อสาร ไร้สาย ทั้งในแบบจำลองและสถานการณ์จริง
- 5.3 การทดสอบและวิเคราะห์ประสิทธิภาพขั้นตอนวิธีที่นำเสนอ

6. ส่วนของงานที่ได้ดำเนินการไปในปี 2548

- 6.1 นำเสนอกรรมวิธีการติดตามใบหน้าและการเคลื่อนที่ของบุคคล โดยใช้กล้องวีดิทัศน์ PTZ 2 ตัว พร้อม โปรแกรมต้นแบบ (รายละเอียดอยู่ในภาคผนวก ก)
- 6.2 ศึกษาและจำลองการส่งสัญญาณวีดิทัศน์ตามมาตรฐาน H.264/MPEG-4 Part 10 ส่งผ่าน ช่องสัญญาณ ไร้สายจำลอง นำเสนอการจัดแผนที่กลุ่มสไลด์และมาโครบล็อก เพื่อกระจายความผิดพลาดที่เกิดขึ้นจากช่องสัญญาณ ไร้สายและได้ข้อสรุปจากการทดลอง (รายละเอียดอยู่ในภาคผนวก ข)
- 6.3 ศึกษาและเสนอโมเดลการควบคุมอัตราใหม่สำหรับมาตรฐาน H.264/MPEG-4 Part 10 และได้ข้อสรุปเบื้องต้นจากการทดลอง (รายละเอียดอยู่ในภาคผนวก ค)

7. ส่วนของงานที่จะดำเนินการต่อไป

- 7.1 พัฒนาโปรแกรมการจัดแผนที่กลุ่มสไลด์และมาโครบล็อกตามสภาพของช่องสัญญาณ และใช้ร่วมกับ เพื่อส่งผ่านช่องสัญญาณ ไร้สายจำลอง
- 7.2 ปรับปรุงประสิทธิภาพของวิธีการตรวจหา ติดตามและรู้จำใบหน้า โดยใช้กล้องวีดิทัศน์แบบสาย กัมเมย และ ชุม 2 ตัวร่วมกันที่นำเสนอ เพื่อให้สามารถประยุกต์ใช้ได้ สถานการณ์จริง
- 7.3 พัฒนาโมเดลการควบคุมอัตราใหม่สำหรับมาตรฐาน H.264/MPEG-4 Part 10 เพื่อการส่งผ่าน ช่องสัญญาณ ไร้สายที่มีเวลาประวิงต่ำ

8. ประโยชน์ของงานวิจัย

- 8.1 พัฒนาองค์ความรู้ด้านการเข้ารหัสสัญญาณวีดิทัศน์เชิงวัตถุและการส่งผ่านโครงข่าย ไร้สาย
- 8.2 พัฒนาระบบจัดระดับปริญญาโทและเอกที่มีทักษะ ความรู้ความสามารถ และมีความรับผิดชอบในหน้าที่
- 8.3 พัฒนาศักยภาพของภาควิชาให้ไปสู่การเป็นผู้นำทางศาสตร์การสื่อสารมัลติมีเดีย

1. จุดประสงค์ของงานวิจัย

- 1.1 พัฒนาระบบและโปรแกรมประยุกต์การติดตามใบหน้าด้วยกล้องวิดิทัศน์ PTZ 2 ตัว ให้ทำงานร่วมกัน เพื่อประยุกต์ใช้ในงานด้านการสอดส่องรักษาความปลอดภัย
- 1.2 พัฒนาระบบให้สามารถประมวลผลตรวจหา ติดตาม และรู้จำใบหน้า ให้ทำงานร่วมกันได้ ในเวลาจริง โดยการเปลี่ยนมุมมองกล้องวิดิทัศน์เพื่อติดตามใบหน้าต้องเป็นไปอย่างอิสระ
- 1.3 สามารถนำประโยชน์ที่ได้จากการใช้กล้องวิดิทัศน์ 2 ตัว ให้ทำงานร่วมกัน มาเพิ่มประสิทธิภาพในกระบวนการรู้จำใบหน้าและระบุบุคคลระหว่างกล้อง
- 1.4 เพื่อเป็นแนวทางในการนำระบบไปประยุกต์ใช้กับลักษณะงานเป้าหมายในสถานการณ์จริงได้

2. งานวิจัยที่ได้ทำไปแล้ว

2.1 ออกแบบกระบวนการตรวจหาใบหน้าโดยใช้ลักษณะสีผิวและลักษณะการเคลื่อนที่ของบุคคลที่สนใจ มาเป็นลักษณะในการตรวจหา ซึ่งทำให้การตรวจหาใบหน้าบุคคลไม่ขึ้นกับพื้นฉากหลัง ส่งผลให้ระบบยังสามารถประมวลผลตรวจหาได้ แม้ว่ากล้องวิดิทัศน์ PTZ จะส่าย ก้มเงย และซูมติดตามบุคคล

- การตรวจหาด้วยสีผิว เลือกใช้การตรวจหาสีผิวด้วยแบบจำลองวงรี ในปริภูมิสี YCbCr โดยทำการสร้างแบบจำลองวงรีที่ใช้ในระบบจากภาพตัวอย่างใบหน้าของกลุ่มบุคคลที่สนใจจำนวน 20 คน

- การตรวจหาการเคลื่อนที่ ใช้กระบวนการเบื้องต้นของวิธีการตรวจหา คือ การหาผลต่างระหว่างเฟรม (frame differencing) ที่ต่อเนื่องกัน เพื่อตรวจหาจุดภาพที่มีการเปลี่ยนแปลงความเข้มสีอันเนื่องจากการเคลื่อนที่ของวัตถุมากกว่าค่าขีดเริ่มเปลี่ยนที่กำหนด

- ออกแบบขั้นตอนการเติมจุดภาพระหว่างจุดภาพการเคลื่อนที่เพื่อให้ได้เป็น โครงร่างของบุคคลที่สนใจเพื่อให้เหมาะสมต่อการนำไปประยุกต์ใช้ร่วมกับการตรวจหาสีผิว

- ออกแบบวิธีการตัดส่วนภาพใบหน้าจากภาพรับเข้าให้มีประสิทธิภาพต่อกระบวนการจับคู่ใบหน้ามากที่สุด โดยพิจารณาจากลักษณะรูปลักษณ์ตามธรรมชาติของรูปทรงใบหน้า จึงเลือกใช้การตัดส่วนภาพใบหน้าในลักษณะวงรี

- ออกแบบขั้นตอนการจำแนกมุมมองของใบหน้าของภาพใบหน้าที่ตรวจหาได้ โดยจำแนกออกเป็น 3 ประเภท คือ ภาพใบหน้าตรง ภาพใบหน้าเอียงด้านซ้าย และ ภาพใบหน้าเอียงด้านขวา เพื่อประโยชน์สำหรับกระบวนการจับคู่ใบหน้า

2.2 ออกแบบกระบวนการรู้จำใบหน้า ซึ่งเมื่อพิจารณาที่ลักษณะของรูปแบบงานวิจัยที่ต้องการ ขั้นตอนวิธีการรู้จำโดยทั่วไปไม่เหมาะต่อการนำมาประยุกต์ใช้กับระบบในเวลาจริง จึงเลือกใช้การจับคู่ใบหน้าระหว่างภาพใบหน้าที่รับเข้ากับภาพใบหน้าในฐานข้อมูลทั้งหมด ว่าภาพใบหน้าในฐานข้อมูลใดที่มีความใกล้เคียงกับภาพใบหน้าที่รับเข้า ก็ระบุว่าเป็นบุคคลนั้น ในการจับคู่ใบหน้าจะใช้วิธีการวัดความคล้าย

(similarity measure) ระหว่างภาพใบหน้า 2 ภาพโดยใช้ลักษณะสี เป็นลักษณะในการวัดความคล้าย ซึ่งใช้ระยะ Bhattacharyya ในการวัดความคล้ายระหว่างภาพ

2.3 ออกแบบลักษณะการทำงานร่วมกันระหว่างกล้องวิดิทัศน์ 2 ตัว โดยแบ่งได้เป็น 2 ลักษณะการทำงาน

- 1 การทำงานร่วมกันในการรู้จำใบหน้า โดยใช้วิธีการจับคู่ใบหน้า นำผลการจับคู่ใบหน้าที่ได้จากกล้องวิดิทัศน์ทั้งสองตัว มาสรุปผลการจับคู่โดยรวมเพื่อเป็นผลการระบุบุคคลของระบบ โดยในการสรุปผลการจับคู่ใบหน้าที่จะดูจากลักษณะมุมมองใบหน้าที่ใช้ในการจับคู่ของกล้องแต่ละตัว ซึ่งจะให้น้ำหนักกับผลการจับคู่ใบหน้าที่ได้จากกล้องวิดิทัศน์ตัวที่สามารถรับภาพมุมมองใบหน้าในมุมมองด้านตรงที่มากกว่า
- 2 การทำงานร่วมกันในการตรวจหาติดตามใบหน้าต่อเนื่องระหว่างกล้องวิดิทัศน์ทั้งสอง กล่าวคือ ระบบจะให้ความสนใจติดตามบุคคลที่สนใจอย่างต่อเนื่อง เพื่อประโยชน์ในการจัดเก็บข้อมูลบุคคลที่สนใจได้อย่างมีประสิทธิภาพในระบบ โดยใช้ลักษณะของสีเสื้อของบุคคลที่สนใจ และลักษณะทิศทางการเคลื่อนที่ของบุคคลที่สนใจ มาเป็นลักษณะสำหรับระบุบุคคลระหว่างกล้องตัวแรกที่ตรวจพบบุคคล ส่งมายังกล้องตัวที่สองในระบบ

3. รายละเอียดของงานวิจัย

งานวิจัยนี้นำเสนอระบบกล้องหลายตัวที่สามารถนำไปประยุกต์ใช้งานสำหรับระบบสอดส่องรักษาความปลอดภัยภายในอาคาร โดยระบบจะประกอบด้วยกล้องวิดิทัศน์ PTZ 2 ตัวซึ่งจะมีการประมวลผลร่วมกันเพื่อประโยชน์ในส่วนการตรวจหาติดตาม และส่วนของการระบุบุคคล โดยใช้วิธีการจับคู่ใบหน้า งานวิจัยนี้จะใช้ประโยชน์ของระบบกล้องหลายตัวและความสามารถของกล้องวิดิทัศน์ PTZ กล่าวคือ ความสามารถในการส่ายและก้มเงยเปลี่ยนมุมมองการรับภาพและการซูมของกล้องวิดิทัศน์ PTZ เพื่อให้ได้ตำแหน่งใบหน้าที่ดีที่สุด ประโยชน์ของระบบกล้องหลายตัว ที่จะมีมุมมองการรับภาพที่กว้างกว่า ทำให้เพิ่มโอกาสในการประมวลผลกับภาพที่มีความเหมาะสมมากที่สุด ประโยชน์ของการส่งต่อพารามิเตอร์ระหว่างกล้องเพื่อช่วยเพิ่มประสิทธิภาพในการประมวลผล

งานวิจัยนี้จึงจะเสนอขั้นตอนวิธีในการทำงานร่วมกันระหว่างกล้องวิดิทัศน์ PTZ ทั้งสอง โดยแบ่งการทำงานออกเป็น 2 ลักษณะ คือ

1. การประมวลผลร่วมกันสำหรับการระบุบุคคล โดยใช้วิธีการจับคู่ภาพใบหน้า การทำงานของทั้งสองกล้องจะทำงานในลักษณะขนาน(parallel) โดยจะประกอบด้วยกระบวนการตรวจหาติดตาม และการจับคู่ใบหน้า ซึ่งจะแบ่งการจับคู่ใบหน้าเป็น 3 กลุ่ม ตามการเรียงตัวของมุมมองใบหน้า (หน้าตรง หน้าเอียงด้านซ้าย หน้าเอียงด้านขวา (โดยระบบจะคัดเลือกภาพใบหน้าที่อยู่ในตำแหน่งที่เหมาะสมกว่า ได้แก่ ใบหน้าตรงหรือค่อนข้างตรง ต่อการนำไปจับคู่ใบหน้าที่ได้จากกล้องวิดิทัศน์ 2 ตัว ซึ่งใน

งานวิจัยนี้เลือกใช้การจับคู่ใบหน้าด้วยลักษณะสีสำหรับจำแนกบุคคล ซึ่งต้องมีข้อมูลภาพใบหน้าอยู่ภายในฐานข้อมูล

2 การทำงานร่วมกันสำหรับการตรวจหาติดตามใบหน้าอย่างต่อเนื่องระหว่างกล้องวีดิทัศน์ 2 ตัว โดยใช้ลักษณะสี แบ่งเป็นส่วนสีเสื้อ ส่วนสีผิว และส่วนสีผม และลักษณะทิศทางการเคลื่อนที่ สำหรับระบบบุคคลระหว่างกล้องวีดิทัศน์ ในกรณีที่บุคคลเป้าหมายเคลื่อนที่พ้นขอบเขตมุมมองของกล้องวีดิทัศน์ตัวแรก กล้องตัวที่สอง จะทำการตรวจหาบุคคลและต้องระบุได้ว่าเป็นบุคคลเป้าหมายเดิม จึงจะทำการถ่าย กัมเมย ติดตามและซุ่มหาเป้าหมาย

สำหรับขั้นตอนวิธีสำหรับตรวจหาใบหน้าที่ต้องประยุกต์ใช้กับกล้องวีดิทัศน์ PTZ ซึ่งเป็นส่วนสำคัญที่ต้องมีการพิจารณาในกรณีที่ต้องทำงานกับกล้องวีดิทัศน์ PTZ งานวิจัยนี้ได้เสนอกระบวนการตรวจหา โดยเลือกใช้ลักษณะสีผิวและลักษณะการเคลื่อนที่ ลักษณะสีผิวเป็นลักษณะหลักในการตรวจหาบริเวณที่มีความน่าจะเป็นส่วนสีผิว และใช้ลักษณะทางการเคลื่อนที่เป็นลักษณะเสริมที่ช่วยเพิ่มประสิทธิภาพในการตรวจหา รวมทั้งลดสัญญาณรบกวนที่เกิดจากบริเวณที่มีสีคล้ายส่วนสีผิว นอกจากนี้ยังได้ใช้กระบวนการประมวลผลภาพเชิงสัญญาณเกี่ยวกับการทำโคเลชันและอีโรชัน กับภาพฐานสองที่ได้จากการตรวจหา เพื่อประโยชน์ในการวิเคราะห์หาตำแหน่งบริเวณใบหน้าที่ต้องการ ซึ่งทำให้ระบบที่นำเสนอสามารถควบคุมการถ่าย และกัมเมยของกล้องวีดิทัศน์ PTZ ได้อย่างเป็นอิสระ และลดความซับซ้อนในการประมวลผล

รูปแบบการทำงานของระบบที่นำเสนอ

รูปแบบการทำงานของระบบที่นำเสนอจะแบ่งออกเป็น 2 รูปแบบ คือ

รูปแบบที่ 1 จะนำเสนอวิธีการประมวลผลจับคู่ใบหน้าเพื่อระบบบุคคลร่วมกันของกล้องวีดิทัศน์ PTZ 2 ตัว เวลาจริง แนวความคิดหลักของรูปแบบการทำงานนี้ คือ ในการประมวลผลสำหรับตรวจหา ติดตาม และจับคู่ใบหน้า บุคคลเคลื่อนที่ ในเวลาจริง ตำแหน่งของภาพใบหน้าที่ถูกถ่ายภาพได้จะมีหลายรูปแบบ บางครั้งอาจรับภาพได้ในแบบหน้าตรง หรือบางครั้งอาจรับภาพได้แบบหน้าเอียง ซึ่งผลกระทบตรงนี้อาจส่งผลไม่มากต่อกระบวนการตรวจหาและติดตาม แต่ส่งผลกระทบต่อประสิทธิภาพในการจับคู่ใบหน้าอย่างมาก เพราะเมื่อพิจารณาจากงานวิจัยที่ผ่านมา [3, 4, 19] ในด้านการรู้จำใบหน้า ระบบจะรู้จำภาพใบหน้าตรงหรือค่อนข้างตรงได้ดีกว่า ด้วยข้อจำกัดดังกล่าว ในการออกแบบระบบจึงแก้ปัญหา โดยตั้งกล้องวีดิทัศน์ PTZ 2 ตัว ในตำแหน่งที่ต่างกัน แต่ต้องมีมุมมองการรับภาพที่เหลื่อมกันอยู่ ทำให้มีความน่าจะเป็นที่กล้องตัวใดตัวหนึ่งในสองตัวจะจับภาพใบหน้าได้ในลักษณะหน้าตรงหรือค่อนข้างตรง ณ เวลาเดียวกัน ซึ่งเป็นข้อได้เปรียบกว่าระบบที่มีกล้องวีดิทัศน์ตัวเดียว เพราะการที่มีกล้องวีดิทัศน์เพียงตัวเดียวอาจมีความเป็นไปได้สูงที่จะจับภาพใบหน้าได้เฉพาะด้านข้าง

รูปแบบที่ 2 นำเสนอลักษณะการทำงานร่วมกันในกระบวนการตรวจหาบุคคลระหว่างกล้องวีดิทัศน์ PTZ 2 ตัว รูปแบบการทำงานจะเป็นการตรวจหาและเคลื่อนที่ติดตามบุคคลเป้าหมาย

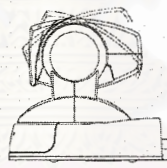
คนเดียวอย่างต่อเนื่องจากกล้องวิดีโอตัวหนึ่งไปยังกล้องวิดีโอตัวที่สอง จะทำให้ระบบสอดส่องและรักษาความปลอดภัยสามารถติดตามการเคลื่อนที่ของบุคคลเป้าหมายที่สนใจได้อย่างต่อเนื่องและมีประสิทธิภาพ กล่าวคือ เมื่อบุคคลเป้าหมายเดินเข้ามาในระบบภายในอาณาบริเวณการรับภาพของกล้องวิดีโอตัวหนึ่ง จะทำการเก็บข้อมูลของบุคคลนั้นในรูปแบบของข้อมูลลักษณะสี จากนั้นเมื่อบุคคลเป้าหมายเคลื่อนที่พ้นจากขอบเขตของกล้องวิดีโอตัวที่หนึ่ง จะส่งข้อมูลนี้ไปยังกล้องวิดีโออีกตัว ทำให้กล้องวิดีโอตัวที่สองนี้จะเปลี่ยนสถานะจากตรวจหาอย่างเดียวเป็นตรวจหาและทำการเปรียบเทียบเพื่อระบุว่าเป็นบุคคลเป้าหมายคนเดิม แนวความคิดหลัก คือ การส่งต่อพารามิเตอร์ระหว่างกล้องวิดีโอ PTZ ทั้งสอง ซึ่งเป็นข้อได้เปรียบกว่ากล้องวิดีโอที่ทำงานเพียงตัวเดียว รวมถึงระบบกล้องหลายตัวที่ไม่มีการใช้ข้อมูลร่วมกันระหว่างกล้อง

กล้องวิดีโอ PTZ

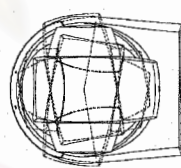
กล้องวิดีโอ PTZ ที่ใช้ในระบบของงานวิจัยนี้ คือ กล้องวิดีโอ SONY รุ่น EVID-100P ดังรูปที่ 1 มีคุณสมบัติที่สำคัญเกี่ยวกับการซูม การก้มเงยและการหมุน ดังนี้



(ก)



(ข)



(ค)

รูปที่ 1 กล้องวิดีโอในการทดลอง(ก)กล้องวิดีโอ EVID-100 (ข)ลักษณะการก้มเงย (ค)ลักษณะการถ่าย

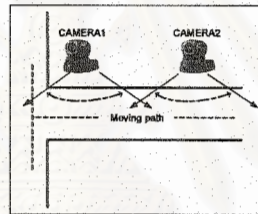
- ในการควบคุมการเคลื่อนที่ของกล้อง ต้องส่งคำสั่ง (command) ควบคุมผ่านทางสายสัญญาณ RS-232
- สามารถซูมเชิงแสง (Optical zoom) ได้สูงสุด 10เท่า
- แนวการถ่ายสูงสุด ± 100 องศา
- แนวการก้มเงยสูงสุด ± 25 องศา
- ความเร็วสูงสุดในการเคลื่อนที่แบบถ่าย 300องศา/วินาที
- ความเร็วสูงสุดในการเคลื่อนที่แบบก้มเงย 125องศา/วินาที
- รองรับการเชื่อมต่อสัญญาณระหว่างกล้องกับส่วนประมวลผลได้ 2รูปแบบ คือ แบบสายสัญญาณภาพรวม (composite video cable) และสายสัญญาณวิดีโอแบบเอสวีดีโอ (S-video cable)

ลักษณะโครงการจัดวางตำแหน่งกล้องวิดีโอ PTZ ทั้ง 2 ตัว

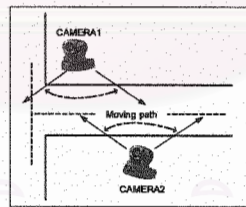
ในหัวข้อนี้จะอธิบายถึงโครงการจัดวางกล้องในแต่ละรูปแบบการทำงาน ดังนี้

รูปแบบที่ 1 (วิธีการประมวลผลสำหรับจับคู่ใบหน้าร่วมกัน) เพื่อให้รองรับแนวความคิดดังที่กล่าวไว้ในตอนต้น การจัดวางกล้องแบบนี้ ต้องเน้นที่มุมมองการรับภาพของทั้งสองกล้องวิดีโอที่ต้องเหลื่อมกัน ในมุมที่เหมาะสมและเกิดประโยชน์ต่อการรับภาพใบหน้ามาประมวลผลมากที่สุด การจัดวางกล้องวิดีโอสามารถทำได้ 2 รูปแบบ ดังแสดงในรูปที่ 2 และ 3 แต่ในงานวิจัยนี้จะเลือกรูปแบบตามรูปที่ 2 เนื่องจากเป็นรูปแบบที่กล้องทั้งสองสามารถรับภาพได้ในตำแหน่งที่เหมาะสมต่อการนำไปประมวลผลจับคู่ สมมติฐานของสถานการณ์ในการจัดวางแบบนี้ ได้แก่ บริเวณจุดแสดงผลงานหรือสินค้าที่ต้องการความปลอดภัย บริเวณจุดติดต่อประชาสัมพันธ์ เป็นต้น ส่วนการจัดวางดังรูป 3 มีข้อบกพร่องเกี่ยวกับมุมมอง เช่น ในกรณีบุคคลหันไปทางกล้องวิดีโอตัวเดียว ทำให้กล้องวิดีโออีกตัวอาจจับภาพได้เป็นส่วนศีรษะด้านหลัง และในกรณีถ้าบุคคลเดินโดยไม่ได้หันเข้ากล้องวิดีโอตัวหนึ่ง ทำให้กล้องวิดีโอทั้งสองรับภาพใบหน้าได้เฉพาะด้านข้าง จึงไม่ได้ภาพใบหน้าตรงเลย

ตำแหน่งการวางกล้องวิดีโอในรูปแบบที่ 1 คือ สูงจากพื้น 1.4 เมตร เท่ากันทั้งสองตัว และ กล้องวิดีโอทั้งสองอยู่ห่างกัน 1.5 เมตร



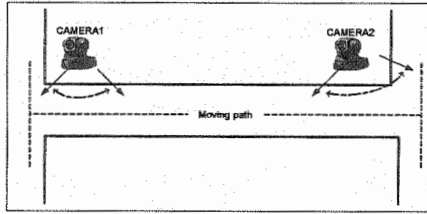
รูปที่ 2 การจัดวางกล้องเพื่อการประมวลผลจับคู่ใบหน้าร่วมกันแบบที่หนึ่ง



รูปที่ 3 การจัดวางกล้องเพื่อการประมวลผลจับคู่ใบหน้าร่วมกันแบบที่สอง

รูปแบบที่ 2 (วิธีการประมวลผลตรวจหาพร้อมกัน) การจัดวางกล้องวิดีโอจะเป็นดังรูปที่ 4 จะเน้นที่กล้องวิดีโอ PTZ ทั้งสองควรจะอยู่ห่างกัน และไม่มีเหลื่อมกัน ในมุมมองการรับภาพ ดังนั้นการจัดวางกล้องจึงไม่ต้องการกำหนดในตำแหน่งมาก (แต่เนื่องจากข้อจำกัดในการทดลอง จึงจำเป็นต้องจัดวางในลักษณะดังรูป) อาจจะวางในมุมตรงข้าม หรืออาจจะอยู่ต่างบริเวณกันก็ได้

ตำแหน่งการวางกล้องวิดีโอในรูปแบบที่ 2 คือ สูงจากพื้น 1.4 เมตร เท่ากันทั้งสองตัว และ กล้องวิดีโอทั้งสองอยู่ห่างกัน 4.9 เมตร



รูปที่ 4 การจัดวางกล้องวิดีโอสำหรับการประมวลผลตรวจหาาร่วมกัน

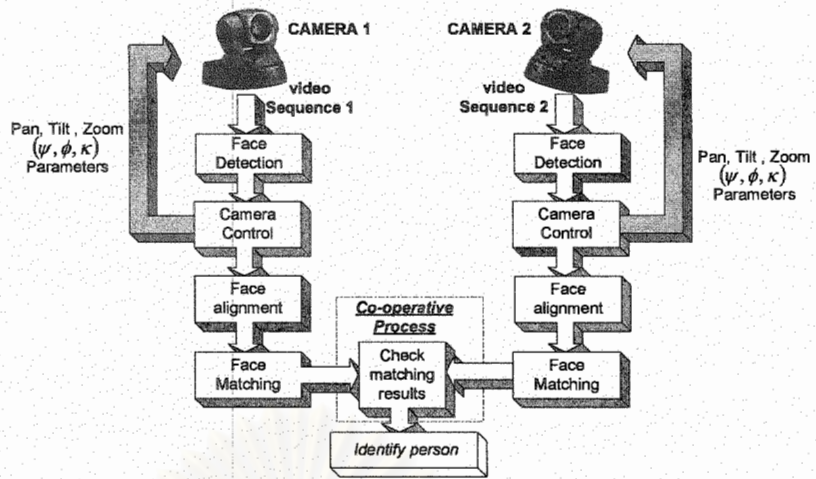
ไคอะแกรมการทำงานโดยรวม

ไคอะแกรมการทำงานของรูปแบบที่ 1 (วิธีการประมวลผลจับคู่ใบหน้าร่วมกัน)

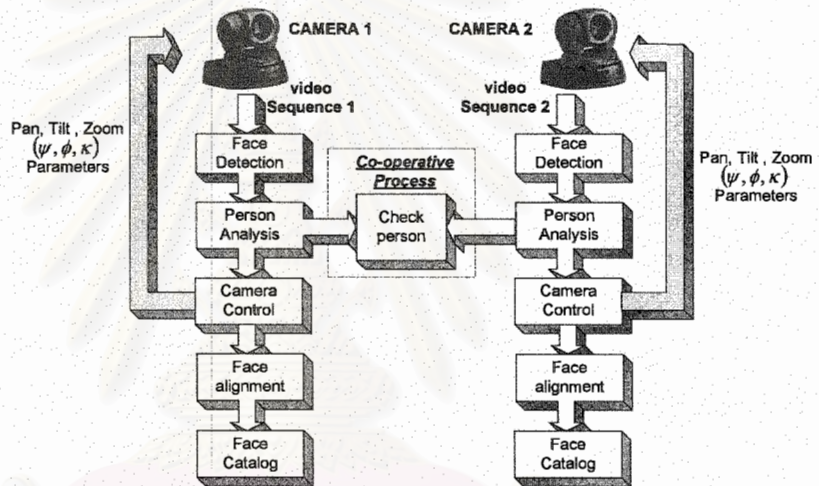
กล้องวิดีโอ PTZ ทั้งสองตัวจะทำงานประมวลผลแบบขนาน (parallel processing) ดังรูปที่ 5 กล่าวคือ กล้องวิดีโอทั้งสองจะประมวลผลในขั้นตอนวิธีที่เหมือนกันและพร้อมกันตลอดเวลา แต่จะมีส่วนการประมวลผลร่วมกันในการวิเคราะห์ผลการจับคู่ใบหน้าที่ได้จากกล้องวิดีโอแต่ละตัวเพื่อสรุปเป็นผลการจับคู่ใบหน้าโดยรวมของระบบ การประมวลผลของกล้องวิดีโอแต่ละตัวจะเริ่มจากการตรวจหาและติดตามใบหน้า ซึ่งจะได้อำนาจตำแหน่งจุดศูนย์กลางของใบหน้าบุคคลที่สนใจ จากนั้นนำตำแหน่งนี้ไปคำนวณหาค่ามุมในการส่ายและก้มเงยของกล้องวิดีโอ ส่วนภาพบริเวณใบหน้าที่ตรวจหาได้ จะนำไปผ่านกระบวนการหามุมมองใบหน้า (face alignment) เพื่อพิจารณาว่าใบหน้าที่รับได้ ณ ขณะนั้นเป็นใบหน้าที่ด้านตรงหรือเอียง สุดท้ายเป็นกระบวนการจับคู่ใบหน้า

ไคอะแกรมการทำงานของรูปแบบที่ 2 (วิธีการประมวลผลตรวจหาาร่วมกัน)

การประมวลผลของกล้องวิดีโอ PTZ จะคล้ายกับกับวิธีการที่กล่าวในส่วนไคอะแกรมการทำงานในรูปแบบวิธีการประมวลผลจับคู่ใบหน้าร่วมกัน ดังรูปที่ 6 แต่จะมีส่วนของการตรวจสอบบุคคลเพิ่มเข้ามาก่อนที่จะมีการส่ายและก้มเงยกล้องวิดีโอติดตามบุคคลเป้าหมาย กล่าวคือ ในกรณีที่กล้องวิดีโอทั้งสองยังตรวจหาไม่พบบุคคล ก็จะทำงานในลักษณะปกติ แต่เมื่อกำลังวิดีโอตัวหนึ่งตรวจหาพบบุคคลแล้ว กล้องวิดีโออีกตัวจะต้องเปลี่ยนสถานภาพการประมวลผล โดยต้องตรวจสอบและระบุว่าเป็นบุคคลคนเดิมเท่านั้น จึงจะสั่งให้กล้องทำการส่ายและก้มเงยติดตาม ขั้นตอนสุดท้ายของการประมวลผล ก็จะทำการเก็บภาพใบหน้าไว้เป็นฐานข้อมูลสำหรับการใช้งานต่อไป



รูปที่ 5 ไคอะแกรมโดยรวมสำหรับวิธีการประมวลผลจับคู่ใบหน้าร่วมกัน



รูปที่ 6 ไคอะแกรมโดยรวมสำหรับวิธีการประมวลผลตรวจหาใบหน้าร่วมกัน

ขั้นตอนวิธีในการตรวจหาและติดตามใบหน้า

ในหัวข้อนี้จะอธิบายถึงรายละเอียดขั้นตอนวิธีในการตรวจหาและติดตามใบหน้าด้วยกล้องวิดีโอ PTZ แนวความคิดหลักของการตรวจหาที่นำเสนอคือ ต้องสามารถประมวลผลตรวจหาได้ในกรณีที่กล้องวิดีโอ PTZ มีการส่ายและก้มเงยเปลี่ยนมุมมองการรับภาพ รวมทั้งแนวการส่ายและก้มเงยของกล้องวิดีโอ PTZ ต้องเป็นไปอย่างอิสระทั้งในสองแนว เมื่อพิจารณาที่ขั้นตอนวิธีที่ผ่านมา [9, 10, 11, 12] ส่วนมากจะต้องมีการเก็บฉากหลังในแต่ละตำแหน่งการส่ายและก้มเงยของกล้อง ทำให้เกิดปัญหาและข้อจำกัด ดังนั้นงานวิจัยนี้จึงได้นำเสนอการประยุกต์ใช้งานร่วมกันระหว่างการตรวจหาด้วยสีผิวและการตรวจหาด้วยการเคลื่อนที่

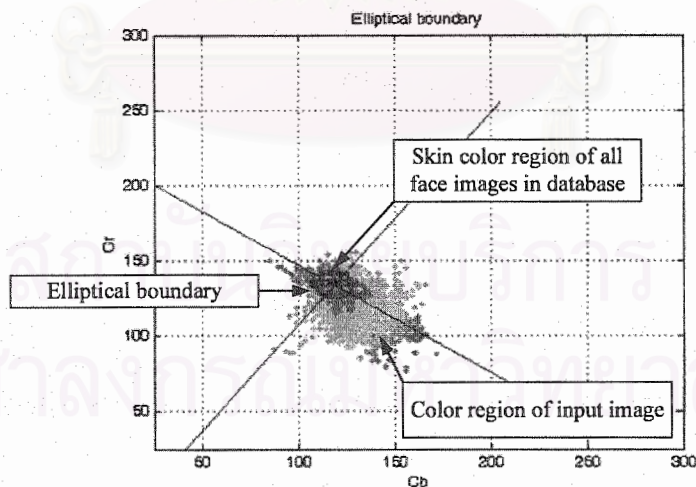
การตรวจหาด้วยสีผิวและการตรวจหาด้วยการเคลื่อนที่สามารถประมวลตรวจหาได้โดยไม่ขึ้นกับฉากหลัง ดังนั้นจึงสามารถประยุกต์ใช้กับกล้องวิดีโอ PTZ ได้อย่างมีประสิทธิภาพ โดยไม่จำกัดตำแหน่งในการถ่ายและกัมเมซของกล้องวิดีโอ และสามารถตรวจหาวัตถุที่มีการเปลี่ยนแปลงรูปร่างได้เหมาะกับการตรวจหาใบหน้าบุคคล

กระบวนการตรวจหาจะเริ่มต้นด้วยการตรวจหาตำแหน่งการเคลื่อนที่ของบุคคล สิ่งที่ได้คือ เค้าโครงโดยรวมของบุคคลที่สนใจ จากนั้นใช้กระบวนการตรวจหาสีผิวเพื่อระบุตำแหน่งของใบหน้าที่อยู่ภายในเค้าโครงบุคคล เหตุผลที่นำการตรวจหาการเคลื่อนที่มาประยุกต์ ใช้ร่วมกับเพื่อลดสัญญาณรบกวนที่เกิดจากวัตถุที่มีสีคล้ายสีผิวในกระบวนการตรวจหาสีผิว รวมทั้งลดบริเวณในการตรวจหาสีผิวให้เหลือแค่บริเวณเค้าโครงบุคคลที่ตรวจหาได้

รายละเอียดในหัวข้อย่อยที่จะกล่าวถึงในหัวข้อนี้ คือ ขั้นตอนวิธีการตรวจหาสีผิวที่ประยุกต์มาจากงานวิจัย [15] กระบวนการตรวจหาการเคลื่อนที่ที่ปรับปรุงจากขั้นตอนวิธีทั่วไปให้เหมาะกับรูปแบบงานที่น่าเสนอ วิธีการกำหนดและหาขอบเขตของใบหน้าเพื่อสามารถนำไปประยุกต์ใช้ในการจับคู่ใบหน้าได้อย่างมีประสิทธิภาพ และวิธีการจำแนกประเภทของมุมมองใบหน้าของภาพที่รับได้

กระบวนการตรวจหาด้วยสีผิว

กระบวนการตรวจหาสีผิวที่เลือกใช้จะอ้างอิงขั้นตอนวิธีจาก [15, 16] ซึ่งจะได้ขอบเขตวงรีที่นำมาใช้ในระบบดังรูปที่ 7



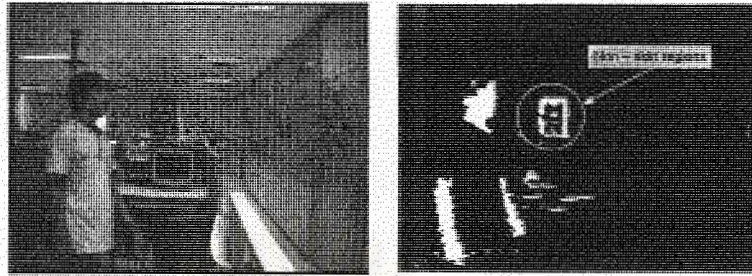
รูปที่ 7 รูปแบบจำลองวงรีที่กำหนดขึ้นจากการฝึกสอน

ปัญหาที่เกิดขึ้นเมื่อใช้แบบจำลองวงรี ในการตรวจหาส่วนสีผิว คือ ปัญหาเกี่ยวกับสิ่งแวดล้อมฉากหลังที่มีสีที่คล้ายกับส่วนสีผิวดังรูปที่ 8 (ส่วนมากวัตถุที่มีผลกระทบจะมีองค์ประกอบสีหลักเป็นสีแดง) ซึ่งทำให้เกิดความยุ่งยากในการวิเคราะห์บริเวณ ทำให้จำเป็นต้องมีขั้นตอนการเลือกขนาด

ของแบบจำลองวงรีที่เหมาะสมต่อการตรวจหาคนที่สุด

รวมทั้งในงานวิจัยนี้จะมีการประยุกต์ใช้การ

ตรวจหาด้วยลักษณะการเคลื่อนที่ร่วมด้วย ซึ่งจะสามารถลดผลกระทบของปัญหาดังได้กล่าวมาได้อย่างมาก



รูปที่ 8 ปัญหาของการตรวจหาสีผิว

กระบวนการตรวจหาการเคลื่อนที่ที่น่าสนใจ

กระบวนการเบื้องต้นของการตรวจหาการเคลื่อนที่โดยทั่วไป คือ การหาความแตกต่างระหว่างเฟรม (frame differencing) ซึ่งจะได้ผลการตรวจหาดังรูปที่ 10 จะเห็นว่าผลที่ได้จะเป็นบริเวณขอบของการเคลื่อนที่และบริเวณบางส่วนภายในขอบเขตการเคลื่อนที่ที่มีค่าความต่างของความเข้มสีอย่างชัดเจน ซึ่งผลที่ได้จะไม่ได้เป็นโครงร่างโดยรวมของบุคคล ทำให้มีความยุ่งยากในขั้นตอนการวิเคราะห์ผลการตรวจหา จึงได้เสนอวิธีการเติมจุดภาพลงไปในพื้นที่ตามแนวนอน ซึ่งต่างจากวิธีเติมเต็มลงในบริเวณที่ถูกล้อมรอบ เนื่องจากในขั้นตอนการตรวจหาการเคลื่อนที่ที่ไม่สามารถรับประกันได้ว่า ผลการตรวจหาการเคลื่อนที่ที่เป็นขอบจะมีการเชื่อมต่อกันตลอดโครงร่างบุคคล ดังรูปที่ 10(ค) ผลการตรวจหาการเคลื่อนที่ในรูปแบบไบนารีแสดงได้ดังสมการที่ (1)

$$x(i, j) = \begin{cases} 1, & \text{motion pixel} \\ 0, & \text{non motion pixel} \end{cases} \quad (1)$$

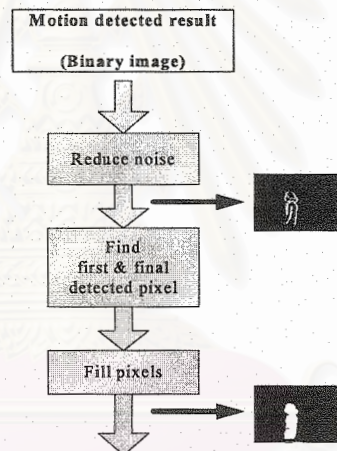
เมื่อ $x(i, j)$ คือ ผลการตรวจหาในรูปแบบไบนารี $i = 0, \dots, \text{Height} - 1$ คือ ตำแหน่งแถวในภาพ $j = 0, \dots, \text{Width} - 1$ คือ ตำแหน่งหลักในภาพ

ไคอะแกรมการทำงานแสดงได้ดังรูปที่ 9 และรูปที่ 10(ง) แสดงตัวอย่างผลการเติมจุดภาพรายละเอียดของการทำงานจะอธิบายได้ดังนี้

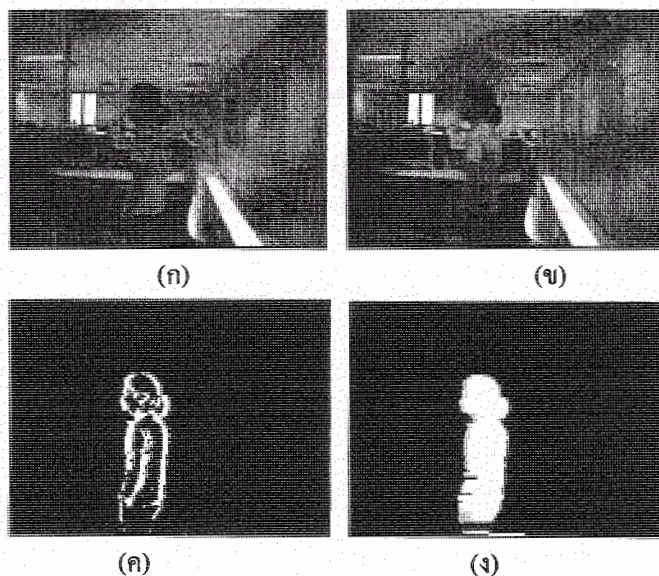
- ภาพไบนารีที่ได้จากการตรวจหาการเคลื่อนที่โดยวิธีทั่วไป จะถูกผ่านขั้นตอนการลดสัญญาณรบกวน
- หาค่าตำแหน่งจุดภาพเริ่มต้นและจุดภาพสุดท้ายในแต่ละแถวของภาพเพื่อใช้เป็นขอบเขตในการเติมจุดภาพ

- ในการหาตำแหน่งจุดภาพเริ่มต้นจะเริ่มพิจารณาตรวจหาจากจุดภาพทางซ้ายสุดของแต่ละแถว จนเมื่อพบจุดภาพตำแหน่งแรกที่ถูกกำหนดว่าเป็นจุดภาพเคลื่อนที่ (motion pixel) จะกำหนดตำแหน่งจุดภาพนี้เป็น ตำแหน่งเริ่มต้นในการเติมจุดภาพ
- ในการหาตำแหน่งจุดภาพสุดท้าย จะเริ่มพิจารณาตรวจหาจากตำแหน่งจุดภาพทางขวาสุดของแต่ละแถว จนเมื่อพบจุดภาพตำแหน่งแรกที่ถูกกำหนดว่าเป็นจุดภาพเคลื่อนที่ (motion pixel) จะกำหนดตำแหน่งจุดภาพนี้เป็น ตำแหน่งสุดท้ายในการเติมจุดภาพ
- ขั้นตอนการเติมจุดภาพจะทำการเปลี่ยนแปลงค่าของภาพในบริเวณของจุดภาพที่อยู่ระหว่างตำแหน่งจุดภาพเริ่มต้นและจุดภาพสุดท้ายจากค่า 0 เป็น 1 หรือ จาก 1 เป็น 1 ดังสมการที่ (2) เมื่อพิจารณาที่แถวภาพที่ i

$$x_i(j) = 1 \quad ; \text{start point} < j < \text{stop point} \quad (2)$$

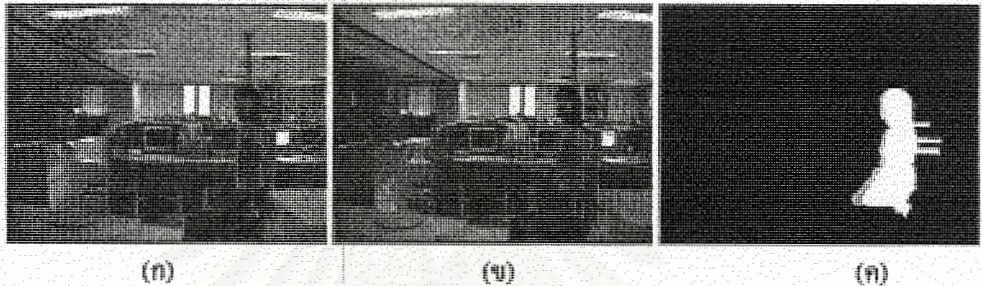


รูปที่ 9 ไคอะแกรมการทำงานในส่วนการต่อเชื่อมจุดภาพ



รูปที่ 10 ผลการตรวจหาการเคลื่อนที่ (ก) เฟรมภาพรับเข้าที่เวลา $t-1$ (ข) เฟรมภาพรับเข้าที่เวลา t (ค) ผลการตรวจหาเบื้องต้นในรูปแบบภาพฐานสอง (ง) ผลการตรวจหาเมื่อผ่านกระบวนการเติมจุดภาพ

ปัญหาของการเติมจุดภาพดังที่กล่าวมา คือ ในกรณีที่ภาพฐานสองของผลการตรวจหาเบื้องต้นไม่เป็นไปตามผลในอุดมคติที่ต้องการ เพราะยังเหลือมีสัญญาณรบกวนที่เกิดจากสิ่งแวดล้อม เช่น ผลกระทบของหน้าจอคอมพิวเตอร์ (ดังรูปที่ 11) เป็นต้น แม้จะผ่านกระบวนการลดสัญญาณรบกวนแล้วก็ตาม จึงต้องมีกระบวนการตรวจเช็คซ้ำหลังจากผ่านกระบวนการเติมจุดภาพ สามารถนำภาพฐานสองนี้ไปวิเคราะห์ต่อไปได้หรือไม่



รูปที่ 11 ผลการเติมจุดภาพที่มีปัญหา(ก)เฟรมภาพรับเข้าที่เวลา $t-1$ (ข) เฟรมภาพรับเข้าที่เวลา t (ค) ผลการตรวจหาเบื้องต้นในรูปแบบภาพฐานสอง(ง)ผลการเติมจุดภาพที่ผิดพลาด

ช่วงห่างระหว่างเฟรมที่จะนำมาตรวจหาการเคลื่อนที่ด้วยกระบวนการลบความแตกต่าง ในงานวิจัยนี้เลือกใช้เฟรมที่ติดกันเพื่อนำมาใช้ประมวลผล (อัตราเฟรมของสัญญาณวิดีโอที่สนักในการประมวลผล คือ 15 เฟรมต่อวินาที) เพราะจะทำให้ได้ขอบเขตของโครงร่างบุคคลที่ใกล้เคียงกับขอบเขตจริง ซึ่งถ้าเลือกช่วงห่างมากเกินไปจะได้ผลการตรวจหาที่มีลักษณะคล้ายการซ้อนเหลื่อมของบุคคล ทำให้ไม่เหมาะสมในการประยุกต์ใช้กับขั้นตอนวิธีการเติมจุดภาพที่ได้กล่าวไว้

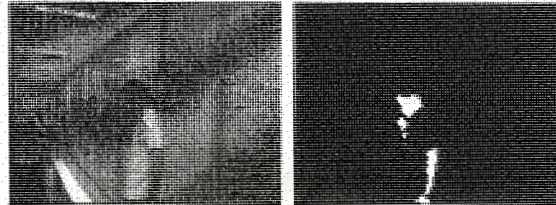
กระบวนการตรวจหาโดยใช้การเคลื่อนที่และสีผิว

กระบวนการตรวจหาโดยรวมของระบบ จะนำการตรวจหาทั้งการตรวจหาด้วยสีผิว และการตรวจหาด้วยการเคลื่อนที่มาทำงานร่วมกัน โดยผลการตรวจหาจะได้ดังสมการที่ (3)

$$x(i, j) = \begin{cases} 1 & ; \text{(motion pixel) and (skin pixel)} \\ 0 & ; \text{otherwise} \end{cases} \quad (3)$$

ลำดับการประมวลผล คือ เริ่มต้นใช้กระบวนการตรวจหาการเคลื่อนที่ เพื่อหาโครงร่างของบุคคลที่เกิดจากการเคลื่อนที่ จากนั้นตรวจหาสีผิวในบริเวณโครงร่างดังกล่าว ซึ่งจะทำให้ลดจำนวนในการประมวลผลต่อจุดภาพและเพิ่มประสิทธิภาพในการตรวจหา โดยจะลดสัญญาณรบกวนที่เกิดจากวัตถุที่มีสีคล้ายสีผิวมนุษย์ ดังตัวอย่างสัญญาณรบกวนรูปที่ 8 และประเด็นสำคัญคือ สามารถนำไปประยุกต์

ตรวจหากับสัญญาณวิดีโอที่รับจากกล้องวิดีโอ PTZ ได้โดยไม่ต้องคำนึงถึงผลกระทบของการเปลี่ยนฉากหลัง รูปที่ 12 แสดงผลลัพธ์ในการตรวจหาเมื่อรวมทั้งสองวิธี



(ก)

(ข)

รูปที่ 12 ตัวอย่างผลการตรวจหาด้วยการตรวจหาการเคลื่อนที่และตรวจหาสีผิว (ก)ภาพรับเข้า(ข)ผลการตรวจหา

การจำแนกเฟรมภาพในการประมวลผล

ในการประมวลผลภาพเพื่อตรวจหาการเคลื่อนที่จากสัญญาณวิดีโอที่รับภาพจากกล้องวิดีโอ PTZ สิ่งหนึ่งที่ต้องพิจารณา คือ ลักษณะของเฟรมที่ขึ้นอยู่กับ การเคลื่อนที่ของกล้อง ซึ่งสามารถแบ่งได้เป็น

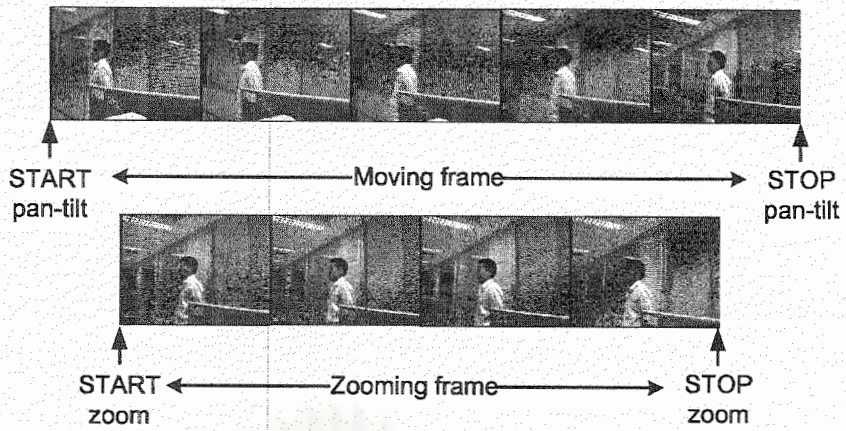
-เฟรมภาพปกติ (normal frame) เป็นเฟรมภาพทั่วไป ที่กล้องวิดีโอไม่มีการส่าย ก้มเงย และการซูม เพราะฉะนั้นสามารถประมวลผลตรวจหาใบหน้าได้ตามปกติ

-เฟรมภาพขณะกล้องวิดีโอ PTZ กำลังเปลี่ยนมุมมองการรับภาพ)เฟรมภาพเคลื่อนที่ (moving frame)(เป็นเฟรมภาพที่ไม่สามารถนำมาวิเคราะห์การตรวจหาได้ เนื่องจากผลกระทบของการส่าย ก้มเงย(แสดงตัวอย่างผลกระทบดังรูปที่ 14) และ การซูม (แสดงตัวอย่างผลกระทบดังรูปที่ 15) ดังนั้นจะไม่มีกรตรวจหาใบหน้าถ้าเป็นเฟรมภาพแบบนี้ รูปที่ 13 จะแสดงให้เห็นถึงขั้นตอนการเปลี่ยนแปลงในลำดับภาพวิดีโอตั้งแต่เฟรมปกติ กล้องวิดีโอเริ่มเปลี่ยนมุมมอง จนเฟรมภาพกลับมาอยู่ในสถานะปกติ

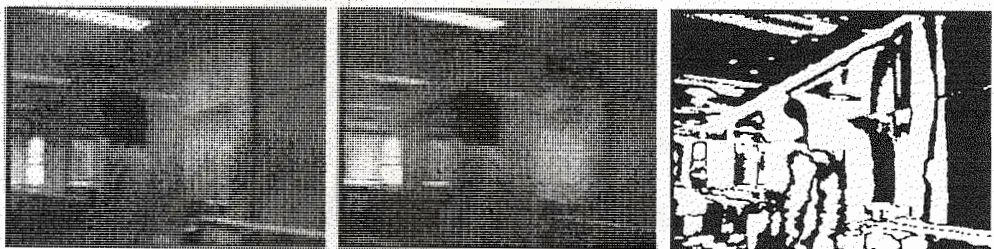
ในการจำแนกเฟรมภาพจะตรวจจากจำนวนจุดภาพเคลื่อนที่ทั้งหมด ดังสมการที่ (4)

$$\text{if } N_{motion} \begin{cases} \geq \tau_f & ;= \text{normal frame} \\ < \tau_f & ;= \text{moving frame} \end{cases} \quad (4)$$

เมื่อ N_{motion} คือ จำนวนจุดภาพทั้งหมดในเฟรมภาพที่ถูกพิจารณาว่าเป็นจุดภาพเคลื่อนที่ และ τ_f คือ ค่าขีดเริ่มเปลี่ยนในการแบ่งเฟรมภาพปกติกับเฟรมภาพเคลื่อนที่



รูปที่ 13 ลำดับวีดิทัศน์ที่เกิดจากการส่าย ก้มเงย และซูมของกล้องวีดิทัศน์ จากหยุดนิ่ง เคลื่อนที่ และหยุดนิ่ง



ก (เฟรม $t-1$)

ข (เฟรม t)

ค (ผลการตรวจหา)

รูปที่ 14 ผลการตรวจหาการเคลื่อนที่ ณ เฟรมภาพที่กล้องกำลังเปลี่ยนมุมมองภาพในลักษณะ ส่ายและก้มเงย



ก (เฟรม $t-1$)

ข (เฟรม t)

ค (ผลการตรวจหา)

รูปที่ 15 ผลการตรวจหาการเคลื่อนที่ ณ เฟรมภาพที่กล้องกำลังเปลี่ยนมุมมองภาพในลักษณะซูม

กระบวนการลดสัญญาณรบกวน

หลังจากผ่านกระบวนการตรวจหา ทั้งการตรวจหาด้วยสีผิวและการตรวจหาการเคลื่อนไหว แม้อัตราสัญญาณรบกวนที่เกิดจากฉากหลังที่มีสีคล้ายกับสีผิว จะถูกตัดทิ้งไป เหลือเพียงแค่วิเวณสีผิวแท้จริง แต่ก็ยังไม่เหมาะสมต่อการนำมาวิเคราะห์ผล เนื่องจากเหตุผล คือ ยังมีสัญญาณรบกวนในระดับจุดภาพเล็กน้อย ซึ่งมีผลต่อการวิเคราะห์บริเวณที่เชื่อมต่อกัน ดังนั้นจึงเลือกใช้ตัวดำเนินการแบบเปิด ในรูปแบบการประมวลผลเชิงสัญญาณ ซึ่งทำกระบวนการอีโรชันแล้วตามด้วยการทำไคเลชันอย่างละหนึ่งรอบ กล่าวคือ หลังจากผ่านกระบวนการ กลุ่มของจุดภาพที่ได้จากการตรวจหาจะมีการเชื่อมต่อกันมากขึ้นและสามารถตัดสัญญาณรบกวน นอกจากนั้นยังช่วยให้การวิเคราะห์บริเวณที่เชื่อมต่อกันมีประสิทธิภาพมากขึ้น ด้วย โดยโครงสร้างส่วนย่อยที่เลือกใช้สำหรับกระบวนการไคเลชันและอีโรชัน จะเป็นดังรูปที่ 16

0	1	0
1	1	1
0	1	0

รูปที่ 16 โครงสร้างส่วนย่อยในการทำอีโรชันและไคเลชัน

การจำแนกส่วนบริเวณใบหน้า

ในการจำแนกบริเวณใบหน้าจากผลการตรวจหา จะพิจารณาที่ผลการตรวจหาด้วยสีผิว ที่ผ่านการลดสัญญาณรบกวนด้วยการตรวจหาด้วยการเคลื่อนที่แล้ว ซึ่งในโครงร่างบุคคล จะประกอบด้วยบริเวณสีผิวที่สามารถตรวจหาได้ 3 ส่วนหลัก คือ ใบหน้า และมือ(รวมแขน)ทั้งสองด้าน ลักษณะที่ใช้แยกส่วนบริเวณใบหน้า que เลือกใช้ในงานวิจัยนี้ คือ ลักษณะทางขนาด ตำแหน่ง และรูปร่าง

- ลักษณะขนาด ใช้กรอบบริเวณที่มีขนาดเล็กมากจนไม่มีความน่าจะเป็นที่จะเป็นบริเวณใบหน้า และบริเวณที่มีขนาดใหญ่จนเกินไป

- ลักษณะตำแหน่ง สมมติบุคคลมีรูปแบบการเคลื่อนที่และการแสดงท่าทางที่เป็นปกติ คือ เดินในท่าทางปกติ จึงสมมติให้ตำแหน่งความน่าจะเป็นของบริเวณใบหน้าที่ตำแหน่งสูงสุดในตัวเลือกกลุ่มบริเวณทั้งหมด

- ลักษณะรูปร่าง รูปร่างโดยทั่วไปของใบหน้าบุคคล จะค่อนข้างกลม จึงเป็นลักษณะหลักที่นำมาใช้ในการจำแนก

ในกรณีพิจารณาการใส่เสื้อของบุคคลเป้าหมายในการตรวจหา

- สวมเสื้อแขนยาว ลักษณะที่จำเป็นในการแยก คือ ขนาดและตำแหน่ง

- สวมเสื้อแขนสั้น ลักษณะทางรูปร่างจะมีความจำเป็นมาก เนื่องจากเมื่อพิจารณาที่รูปร่างท่อนแขน จะมีลักษณะคล้ายทรงกระบอก ซึ่งจะต่างกับใบหน้าที่เป็นทรงกลมอย่างชัดเจน

การกำหนดขอบเขตให้กับบริเวณใบหน้า

เป้าหมายหลักของการนำไปประยุกต์ใช้ของระบบในงานวิจัยนี้ คือ การตรวจหาและติดตามใบหน้าเพื่อนำภาพใบหน้าที่ได้ไปใช้สำหรับการจับคู่ใบหน้า ในสิ่งแวดล้อมจริง ซึ่งมีปัจจัยที่สำคัญที่ควรนำมาพิจารณา 2 ประการ คือ ผลกระทบของฉากหลังในบริเวณใกล้เคียงกับใบหน้าที่จะติดอยู่ในภาพใบหน้า และตำแหน่งมุมมองของใบหน้า

โดยทั่วไปงานทางด้านความรู้จำภาพใบหน้า [3, 4, 19] จะใช้การตัดส่วนภาพ (cropping) ใบหน้าเป็นลักษณะภาพสี่เหลี่ยมจากภาพรับเข้า และนำภาพใบหน้าที่ได้นี้ไปใช้สำหรับกระบวนการรู้จำเลย และส่วนใหญ่จะกำหนดลักษณะพื้นฉากหลังเป็นพื้นฉากหลังสีเรียบ ดังนั้นผลกระทบของฉากหลังจึงไม่ส่งผลกระทบต่อประสิทธิภาพการรู้จำใบหน้ามากนัก แต่ในงานวิจัยนี้ได้ทำการทดลองในห้องปฏิบัติการวิจัยกรรมวิธีสัญญาณดิจิทัลที่ไม่ได้กำหนดลักษณะฉากหลัง รวมทั้งกล้องวิดีโอที่สนับสนุนการรับภาพเพื่อติดตามบุคคลทำให้ฉากหลังเปลี่ยนแปลงไป ดังนั้นในกรณีระบบที่นำเสนอ ถ้าใช้การตัดส่วนภาพใบหน้าในลักษณะสี่เหลี่ยมเพื่อนำมาใช้ในการจับคู่ใบหน้า จะส่งผลกระทบต่อประสิทธิภาพการจับคู่ใบหน้า เพราะฉากหลังของภาพใบหน้าที่เก็บไว้ในฐานข้อมูลจะต่างจากฉากหลังที่รับภาพได้ในขณะนั้น จึงต้องมีการตัดส่วนภาพใบหน้าในลักษณะทางเรขาคณิตรูปแบบอื่นเพื่อลดผลกระทบของฉากหลังให้น้อยที่สุด เมื่อพิจารณาที่โครงสร้างของใบหน้าบุคคลโดยทั่วไป จะเห็นว่ารูปทรงจะคล้ายรูปวงรีหรือวงกลม จึงมีความเหมาะสมที่จะประยุกต์ใช้การตัดส่วนภาพใบหน้าในลักษณะภาพวงรีหรือวงกลมมากที่สุด

ดังนั้นงานวิจัยนี้จึงได้เลือกลักษณะการตัดส่วนภาพใบหน้าในลักษณะวงรีหรือวงกลม ขั้นตอนการตัดส่วนภาพจะแบ่งออกเป็น 2 ขั้นตอน คือ ขั้นที่ 1 การกำหนดขอบเขตอย่างคร่าวๆของบริเวณใบหน้าในลักษณะสี่เหลี่ยม และขั้นที่ 2 การตัดส่วนภาพในลักษณะวงรีหรือวงกลม พิจารณาจากภาพผลการตรวจหาในอุดมคติในรูปที่ 17 เมื่อผลกระบวนการทำป้าย (labeling) จะได้รับเวกเตอร์สีผิวทั้งหมด 3 กลุ่ม ขั้นตอนแรกในการกำหนดขอบเขตแบบสี่เหลี่ยม จะกำหนดได้ดังนี้

ขอบบน เลือกจากตำแหน่งสูงสุดของผลการตรวจหา

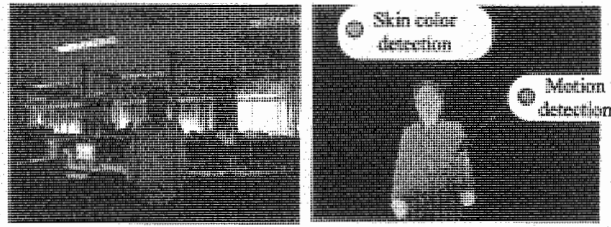
ขอบล่าง เลือกจากตำแหน่งจุดภาพต่ำสุดของบริเวณกลุ่มใบหน้าจากผลการตรวจหาสีผิว

ขอบซ้าย เลือกจากตำแหน่งจุดภาพที่อยู่ตำแหน่งซ้ายสุด โดยเลือกเฉพาะแถวภาพระหว่าง

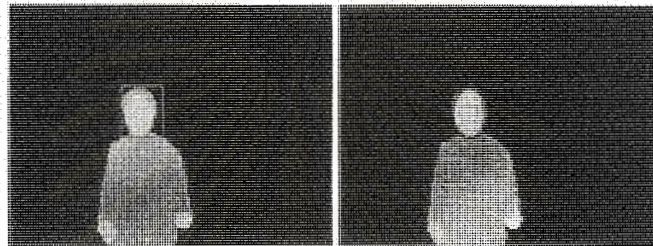
ขอบบนกับขอบล่าง เท่านั้น

ขอบขวา เลือกจากตำแหน่งจุดภาพที่อยู่ตำแหน่งขวาสุด โดยเลือกเฉพาะแถวภาพระหว่าง

ขอบบนกับขอบล่าง เท่านั้น



รูปที่ 17 ผลการตรวจหาในอุดมคติ



(ก)

(ข)

รูปที่ 18 ขั้นตอนการกำหนดขอบเขตใบหน้า (ก) กำหนดขอบเขตในลักษณะสีเหลือง(ข) กำหนดขอบเขตในลักษณะวงรี

จะเห็นว่าการกำหนดขอบเขตสีเหลืองเป็นการกำหนดบริเวณอย่างคร่าว ๆ ดังตัวอย่างรูปที่ 18(ก) เพื่อใช้เป็นบริเวณสำหรับการหาขอบเขตของวงรี ดังตัวอย่างรูปที่ 18(ข)และตำแหน่งที่เหมาะสมที่สุด (position optimization) การหาขนาดของวงรีจะหาได้ดังนี้

ขนาดแกนเอก(แกนตั้ง)หาจากขนาดความสูงของขอบเขตสีเหลือง

ขนาดแกนโท(แกนนอน)หาจากขอบเขตที่อยู่ทางซ้ายสุดและทางขวาสุดของกลุ่มบริเวณที่ตรวจหาได้ แต่ต้องอยู่ภายในขอบเขตสีเหลืองที่ได้จากขั้นตอนแรก โดยจะตัดแถวที่มีความยาวมากกว่าและน้อยกว่าค่าขีดเริ่มเปลี่ยนในการตัดสินใจที่กำหนดไว้เมื่อเปรียบเทียบกับขนาดความยาวโดยเฉลี่ย

การจำแนกประเภทใบหน้าตามลักษณะมุมมอง (Face alignment)

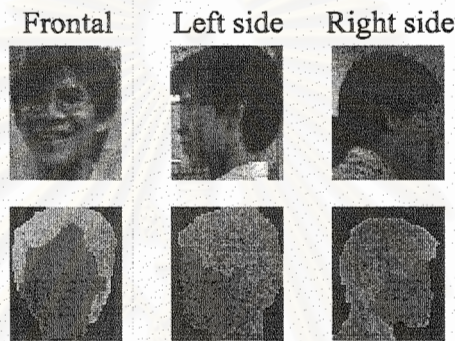
เมื่อกำหนดขอบเขตสำหรับการตัดส่วนภาพใบหน้าแล้ว สิ่งที่ต้องวิเคราะห์ คือ เป็นภาพใบหน้าในมุมมองใด ซึ่งไม่จำกัดเฉพาะใบหน้าตรงเท่านั้น เหตุผลที่ต้องมีการวิเคราะห์ในหัวข้อนี้ เนื่องจากเป็นระบบสำหรับนำภาพใบหน้า ไปใช้จับคู่ใบหน้า มุมมองของใบหน้ามีส่วนสำคัญอย่างมากในประสิทธิภาพของการจับคู่

โดยงานวิจัยนี้ได้จำแนกประเภทมุมมองของใบหน้าเป็น 3ประเภท คือ .1ใบหน้าตรง .2 ใบหน้าด้านข้างซ้าย .3 ใบหน้าด้านข้างขวา ดังรูปที่ 19วิธีการจำแนกจะวิเคราะห์จากผลการตรวจหาสีผิวเทียบกับผลการตรวจหาการเคลื่อนที่ ดังนี้

กำหนดให้กลุ่มสีผิวที่ตรวจหาได้ บริเวณสีแดงในรูปที่ 19 มีตำแหน่งจุดศูนย์กลาง (center of gravity) ที่แถว i_s และคอลัมน์ที่ j_s และให้บริเวณที่ตรวจหาได้จากการตรวจหาสีผิวและการเคลื่อนที่ บริเวณสีแดงรวมกับสีเขียว เฉพาะขอบเขตวงรีบริเวณใบหน้า มีตำแหน่งจุดศูนย์กลางที่แถว i_h และคอลัมน์ที่ j_h จะสามารถจำแนกประเภทใบหน้าตามลักษณะมุมมอง ได้เป็น

- กลุ่มใบหน้าตรง เมื่อ $|j_s - j_h| < \tau_{front}$
- กลุ่มใบหน้าเอียงด้านซ้ายซ้าย เมื่อ $j_s - j_h < \tau_{left}$
- กลุ่มใบหน้าเอียงด้านซ้ายขวา เมื่อ $j_s - j_h < \tau_{right}$

เมื่อ τ_{front} คือ ค่าขีดเริ่มเปลี่ยนของการตัดสินใจเป็นลักษณะใบหน้าตรง τ_{left} คือ ค่าขีดเริ่มเปลี่ยนของการตัดสินใจเป็นลักษณะใบหน้าเอียงด้านซ้ายมีค่าติดลบ τ_{right} คือ ค่าขีดเริ่มเปลี่ยนของการตัดสินใจเป็นลักษณะใบหน้าเอียงด้านขวามีค่าเป็นบวก



รูปที่ 19 ลักษณะการจำแนกใบหน้าเป็น 3 ประเภท

การควบคุมการส่าย ก้มเงย และขมของกล้องวิดีโอ PTZ เพื่อติดตามใบหน้าบุคคลเป้าหมาย

เป็นกระบวนการสำหรับคำนวณหามุมและอัตราการขมในการเปลี่ยนมุมมองของกล้องวิดีโอ โดยเป้าหมาย คือ พยายามเปลี่ยนมุมมองการรับภาพของกล้องให้ใบหน้าบุคคลเป้าหมายมาอยู่บริเวณกึ่งกลางภาพ ขั้นตอนการคำนวณพารามิเตอร์ต่าง ๆ จะเลือกใช้เฉพาะข้อมูลตำแหน่งจุดภาพในพิกัด 2 มิติเท่านั้น โดยสมมติว่ามุมในการเปลี่ยนมุมมองแต่ละขั้น (step) ไม่มาก ซึ่งประมาณได้ว่าการคำนวณหามุมด้วยพิกัดตำแหน่งใน 2 มิติ จะมีค่าความผิดพลาดไม่มากเมื่อเปรียบเทียบกับการคำนวณด้วยพิกัดตำแหน่ง 3 มิติ

การคำนวณมุมในการเปลี่ยนมุมมองและอัตราการขม

กำหนดให้ ψ คือ มุมส่าย ϕ คือ มุมก้มเงย และ x คือ ตัวประกอบการขม ดังนั้นสามารถหาค่าพารามิเตอร์ทั้งสามได้ ดังนี้

- มุมส่าย (Pan angle)

$$\psi = \frac{Diff_x}{\beta_x} = \frac{x_f - x_{center}}{\beta_x} \quad (3.5)$$

- มุมก้มเงย (Tilt angle)

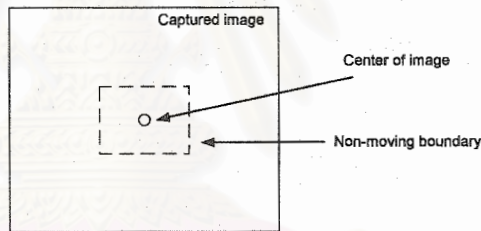
$$\phi = \frac{Diff_y}{\beta_y} = \frac{y_f - y_{center}}{\beta_y} \quad (3.6)$$

-ตัวประกอบการซูม (Zoom factor)

$$\kappa = \frac{Q_{target}}{Q_{detected}} \quad (3.7)$$

เมื่อ $Diff_x$ และ $Diff_y$ คือ ระยะห่างในแนวนอนและแนวตั้งระหว่างตำแหน่งจุดศูนย์กลางของใบหน้า (x_f, y_f) กับตำแหน่งจุดศูนย์กลางภาพ (x_{center}, y_{center}) β_x และ β_y คือ อัตราของจำนวนจุดภาพต่อการเคลื่อนที่ของกล้องไป 1 องศา Q_{target} คือ ขนาดของบริเวณใบหน้าเป้าหมาย และ $Q_{detected}$ คือ ขนาดของบริเวณใบหน้าที่ตรวจหาได้ในหน่วยจุดภาพ

ในกรณีที่ใบหน้าเป้าหมายอยู่บริเวณใกล้กับกึ่งกลางภาพอยู่แล้ว ไม่จำเป็นต้องสั่งให้กล้องวิดิทัศน์เปลี่ยนมุมมอง ดังนั้นจึงได้กำหนดช่วงขอบเขตในภาพเพื่อใช้เป็นขอบเขตในการเคลื่อนที่ดังรูปที่ 20 ถ้าตรวจหาได้ว่าใบหน้ายังอยู่ในช่วงที่กำหนดที่ไม่ต้องมีการเปลี่ยนมุมมอง (Non-moving boundary) ก็จะไม่สั่งกล้องให้เคลื่อนที่ จนกว่าใบหน้าเป้าหมายจะอยู่นอกขอบเขต ก็จะสั่งกล้องให้เคลื่อนที่



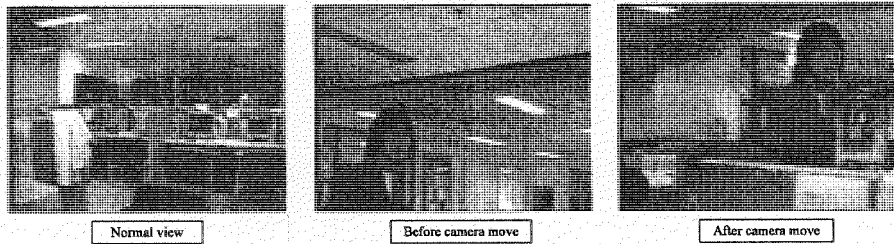
รูปที่ 20 ขอบเขตการสั่งให้กล้องเคลื่อนที่

การชดเชยผลกระทบเนื่องจากการซูม

ในกระบวนการคำนวณพารามิเตอร์การส่ายและก้มเงย เมื่อมุมมองการรับภาพถูกซูมขยายแล้ว จะเกิดปัญหาขึ้น ดังตัวอย่างปัญหาในรูปที่ 21 เนื่องจากเมื่อก้องมีการซูมตำแหน่งของจุดภาพจะถูกสเกลด้วยอัตราส่วนการซูมดังสมการที่ (8) ดังนั้นตำแหน่ง (x_f, y_f) ที่ใช้คำนวณในสมการที่ (5) และ (6) จะต้องถูกสเกลกลับมายังพิกัดปกติก่อน โดยคำนวณย้อนกลับจากสมการที่ (8)

$$X_z = \frac{f'}{f} \begin{bmatrix} x_n \\ y_n \end{bmatrix} \quad (8)$$

เมื่อ X_z คือ ตำแหน่งจุดภาพที่ถูกสเกลด้วยตัวประกอบ f'/f (x_n, y_n) คือ ตำแหน่งจุดภาพ ณ มุมมองปกติ f' คือ ความยาวโฟกัสหลังการซูม และ f คือ ความยาวโฟกัสก่อนการซูม



รูปที่ 21 ปัญหาของการเปลี่ยนมุมมองกล้องวิดีโออัตโนมัติเมื่อมีการหมุน

ขั้นตอนวิธีการเปรียบเทียบใบหน้ากับใบหน้ารวมในฐานข้อมูล

กระบวนการวัดความคล้าย

ในกระบวนการวัดความคล้ายจะประยุกต์ใช้การวัดความคล้ายด้วยระยะ Bhattacharyya มาใช้สำหรับเป็นค่าตัดสินในการเปรียบเทียบภาพใบหน้าที่รับภาพและตัดส่วนภาพได้กับภาพใบหน้าที่ถูกเก็บไว้ในฐานข้อมูล เนื่องจากในระบบที่นำเสนอ กล้องวิดีโอ PTZ จะเปลี่ยนมุมมองการรับภาพตลอดเวลา ทำให้ภาพใบหน้าที่รับได้มีหลายมุมมอง รวมทั้งมีการสเกลที่เกิดจากการซูมทำให้ขนาดภาพใบหน้าที่มีขนาดเปลี่ยนแปลงอยู่เสมอ ซึ่งเป็นข้อดีของการเปรียบเทียบโดยใช้ระยะ Bhattacharyya เพราะสามารถเปรียบเทียบระหว่างชุดข้อมูลที่มีขนาดไม่เท่ากัน รวมทั้งมีงานวิจัยต่าง ๆ [17, 18, 24] ที่นำระยะนี้ไปประยุกต์ใช้ ในลักษณะการประยุกต์ที่ใกล้เคียงกับระบบที่นำเสนอ

โดยในการเปรียบเทียบภาพใบหน้า จะใช้ข้อมูลค่าความเข้มสีของภาพใบหน้าเฉพาะตำแหน่งจุดภาพที่อยู่ในบริเวณขอบเขตวงรี หรือวงกลม เท่านั้น เนื่องจากต้องการลดผลกระทบของฉากหลัง ที่จะส่งผลกระทบต่อประสิทธิภาพในการจับคู่ใบหน้าที่ถูกต้อง

การเปรียบเทียบฮิสโทแกรมด้วยระยะ Bhattacharyya [17, 23]

ระยะ Bhattacharyya เป็นรูปแบบการเปรียบเทียบลักษณะหนึ่งในกระบวนการวัดความคล้าย ระหว่างชุดการกระจายตัวของข้อมูล 2 ชุด ซึ่งได้ถูกนำมาประยุกต์ใช้ในกระบวนการทางการประมวลผลภาพสำหรับการเปรียบเทียบฮิสโทแกรมของสองกลุ่มข้อมูลที่ไม่จำเป็นต้องเป็นค่าฮิสโทแกรมของทั้งภาพ อาจเป็นแค่บริเวณใดบริเวณหนึ่งที่สนใจเท่านั้น โดยไม่จำเป็นต้องทั้งสองบริเวณจะต้องมีขนาดเท่ากัน (ในหน่วยจุดภาพ) จึงเป็นข้อดีที่สามารถรองรับ การเปรียบเทียบวัตถุที่มีการเปลี่ยนแปลงรูปร่างได้ ทั้งในลักษณะการเคลื่อนที่เปลี่ยนมุมมองและการเปลี่ยนแปลงขนาด

สมมติให้ \hat{p} เป็นฮิสโทแกรมสีของแบบจำลองเป้าหมายและ \hat{q} เป็นฮิสโทแกรมสีของแบบจำลองรับเข้าที่ต้องการนำมาเปรียบเทียบ 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 เป็นจำนวนถึงสีของฮิสโทแกรม) bin histogram) เปรียบเทียบ

เช่น ถ้าในกรณีเลือกใช้ขนาดขององค์ประกอบสี่แต่ละองค์ประกอบมีขนาด 8บิต ฉะนั้นจะได้ขนาดดั่งสี่เท่ากับ 256ดั่ง ดังนั้นสามารถคำนวณหาสัมประสิทธิ์ Bhattacharyya ได้ดั่งสมการที่ 9)

$$\hat{\rho} = \sum_{u=1}^m \sqrt{\hat{p}_u \times \hat{q}_u} \quad (9)$$

อธิบายความหมายทางเรขาคณิตตามสมการที่ (9) จะแสดงถึงค่าโคไซน์ (cosine) ของมุมระหว่างค่าดั่งสี่ทั้ง m มิติในรูปเวกเตอร์หนึ่งหน่วย $(\sqrt{\hat{p}_1}, \dots, \sqrt{\hat{p}_m})^T$ และ $(\sqrt{\hat{q}_1}, \dots, \sqrt{\hat{q}_m})^T$

จากสมการที่ (9) จะสามารถหาค่าระยะ Bhattacharyya ที่เป็นระยะระหว่างสองกลุ่มของการกระจายตัวของข้อมูล (d_{Bh}) ดั่งสมการที่ (10)

$$d_{Bh} = \sqrt{1 - \rho[\hat{p}, \hat{q}]} \quad (10)$$

การทำงานร่วมกันระหว่างกล้อง

เป็นส่วนการประมวลผลที่เพิ่มขึ้นจากระบบกล้องวิดิทัศน์ตัวเดียว ซึ่งเป็นข้อได้เปรียบของระบบกล้องหลายตัว โดยในงานวิจัยนี้ได้เสนอรูปแบบการทำงานร่วมกันระหว่างกล้องวิดิทัศน์ PTZ 2 ตัว ใน 2 รูปแบบ ดังนี้

การประมวลผลร่วมกันสำหรับระบุบุคคลโดยการจับคู่ใบหน้า

เป็นการทำงานร่วมกันในลักษณะประมวลผลร่วมกัน เพื่อสรุปผลการจับคู่ใบหน้า โดยใช้ผลการจับคู่ที่ได้จากภาพใบหน้าจากทั้งสองกล้องที่ต้องรับภาพใบหน้าในคนละมุมมอง ซึ่งสมมติฐานเบื้องต้นของมุมมองการรับภาพทั้งสองกล้อง คือ ควรจะมีกล้องวิดิทัศน์ตัวใดตัวหนึ่งที่สามารถจับภาพใบหน้าได้ในตำแหน่งหน้าตรงหรือ ก่อนข้างตรง และเมื่อพิจารณาที่งานวิจัยที่ผ่านมาทางด้านการรู้จำ [19] ตำแหน่งใบหน้าที่มีประสิทธิภาพในการจับคู่ใบหน้านั้นมากกว่า จะเป็นมุมมองใบหน้าแนวตรง มากกว่าในมุมมองอื่น ดังนั้นจึงจะให้น้ำหนักในการสรุปผลจากกล้องวิดิทัศน์ PTZ ตัวที่สามารถรับภาพใบหน้าด้านตรงมากกว่ากล้องวิดิทัศน์อีกตัวหนึ่ง

กำหนดให้คะแนนการจับคู่ใบหน้าระหว่างภาพใบหน้ารับเข้ากับภาพใบหน้าในฐานข้อมูลทั้งหมด N ภาพใบหน้าจากบุคคลต้นแบบ M คน ของกล้องแต่ละตัวมีค่าเป็น γ_i เมื่อ $i=1, \dots, N$ เป็นค่าอันดับความคล้ายของภาพใบหน้าเรียงลำดับจากค่าที่มีความคล้ายมากที่สุดไล่ลง โดยค่า γ_i หาได้จากสมการที่ (9) ในการพิจารณาผลของการจับคู่ใบหน้านั้นจะไม่พิจารณาเฉพาะค่าคะแนนการจับคู่ที่สูงสุดเท่านั้น แต่จะพิจารณาจากอันดับความคล้าย P ลำดับสูงสุด ซึ่งใน P ลำดับนี้จะสามารถคำนวณค่าคะแนนการจับคู่ใบหน้าที่ของแต่ละบุคคลต้นแบบได้ดั่งสมการที่ (11)

$$\mu^i = \frac{1}{n^i} \sum_{j=1}^P \gamma_j^i ; i=1, \dots, M \quad (11)$$

เมื่อ μ^i คือ ค่าเฉลี่ยของคะแนนการจับคู่ใบหน้าทีที่ระบุเป็นบุคคลต้นแบบ i ในฐานข้อมูล n^i คือ จำนวนความถี่ใน P ครั้งที่ระบบระบุว่าเป็นบุคคลต้นแบบ i ฉะนั้นจะได้คะแนนการจับคู่ใบหน้าในการระบุบุคคลแต่ละบุคคลในฐานข้อมูลต้นแบบ ID^i ดังสมการที่ 12

$$ID^i = \frac{n^i}{P} \mu^i \quad (12)$$

การที่ไม่ใช้ค่า μ^i ซึ่งเป็นค่าเฉลี่ยของคะแนนโดยตรง ก็เนื่องมาจากต้องมีการพิจารณาความได้เปรียบในเรื่องความถี่ในการระบุบุคคล เช่น ในกรณีมีบุคคลต้นแบบในฐานข้อมูล 5 คน คือ บุคคล ก บุคคล ข บุคคล ค บุคคล ง บุคคล จ พิจารณาใน 20 อันดับสูงสุด ระบุเป็นบุคคล ก 10 ครั้ง บุคคล ข 3 ครั้ง บุคคล ค 1 ครั้ง บุคคล ง 4 ครั้ง และบุคคล จ 2 ครั้ง ดังนั้นจะต้องให้ค่าความได้เปรียบกับบุคคล ก มากที่สุด

โดยทั้งสองกล้องวิดิทัศน์จะประมวลผลในรูปแบบเดียวกันดังที่กล่าวมา ดังนั้นสามารถสรุปผลของการจับคู่ใบหน้าเพื่อระบุบุคคลโดยรวมจากทั้งสองกล้องได้ดังสมการที่ 13

$$ID_{final}^i = \eta \cdot (ID_1^i) + (1 - \eta) \cdot (ID_2^i) \quad (13)$$

เมื่อ ID_{final}^i คือ คะแนนการจับคู่ใบหน้าทีที่ระบุเป็นบุคคลต้นแบบที่ i โดยสรุปผลมาจากกล้องทั้งสองตัว ID_1^i และ ID_2^i คือ คะแนนการจับคู่ใบหน้าทีที่ระบุเป็นบุคคลต้นแบบที่ i จากกล้องตัวที่ 1 และตัวที่ 2 ตามลำดับ ส่วนค่า η เป็นค่าถ่วงน้ำหนักระหว่างกล้องทั้งสองตัว โดย

$\eta = 0.75$ เมื่อกำลังตัวใดตัวหนึ่งสามารถรับภาพใบหน้าได้ในมุมมองหน้าตรง และอีกกล้องหนึ่งเป็นภาพใบหน้ามุมมองด้านข้างให้น้ำหนักกับภาพใบหน้าตรงมากกว่า

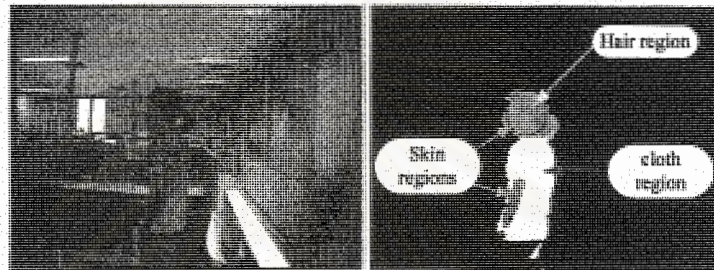
$\eta = 0.50$ เมื่อกำลังสองตัวรับภาพใบหน้าได้ในมุมมองเดียวกัน อาจจะเป็นหน้าตรงหรือหน้ามุมมองด้านข้างทั้งสองให้น้ำหนักเท่ากัน

การทำงานร่วมกันสำหรับตรวจหาติดตามใบหน้าอย่างต่อเนื่อง

เป็นการทำงานร่วมกัน ในลักษณะการตรวจสอบผลการตรวจหาระหว่างกล้องวิดิทัศน์เพื่อระบุความเป็นบุคคลเดียวกัน ซึ่งมีเป้าหมาย คือ การตรวจหาและติดตามอย่างต่อเนื่องระหว่างกล้องวิดิทัศน์ PTZ 2 ตัว เช่น ในสถานการณ์เมื่อกำลังวิดิทัศน์ตัวแรกสามารถตรวจหาพบบุคคล ก็จะเก็บข้อมูลลักษณะ (feature data) ของบุคคลนี้ไว้ ซึ่งข้อมูลชุดนี้จะถูกส่งไปให้กับกล้องวิดิทัศน์ตัวที่สอง กล้องวิดิทัศน์ตัวที่สองนี้จะนำข้อมูลนี้เพื่อใช้เปรียบเทียบบุคคลที่เข้ามาในมุมมองกล้องว่าเป็นบุคคลเดียวกันหรือไม่ ถ้าใช่ก็จะทำการตรวจหาและทำการถ่าย ก้มเงย และซูม กล้องวิดิทัศน์เพื่อติดตามบุคคลต่อเนื่อง เพื่อสามารถจับเก็บภาพบุคคลที่สนใจได้อย่างต่อเนื่องจากกล้องวิดิทัศน์ตัวหนึ่งไปยังอีกตัวหนึ่ง โดยในงานวิจัยนี้เลือกใช้ลักษณะสีและลักษณะทิศทางการเคลื่อนที่ของบุคคลเป็นลักษณะในการใช้เปรียบเทียบบุคคล สำหรับลักษณะสีจะแบ่งออกเป็น 3 ส่วนหลักจากผลการตรวจหาได้ดังนี้

การแบ่งบริเวณบุคคลในส่วนบริเวณที่ตรวจหา

ข้อมูลลักษณะสีที่เลือกใช้ คือ ฮิสโทแกรมของสีในปริภูมิ RGB แล้วใช้การเปรียบเทียบกับระยะ Bhattacharyya ดังรายละเอียดที่กล่าวไว้ในบทที่ 2 โดยจากผลการตรวจหาที่ได้เป็นโครงร่างบุคคลนั้น ถ้าเลือกใช้เฉพาะข้อมูลสีบริเวณใบหน้า จะทำให้ ดังนั้นจะเห็นว่าถ้านำข้อมูลสีบริเวณเสื้อผ้าของบุคคลมาใช้ในการเปรียบเทียบกับ ก็น่าจะเพิ่มค่าความมั่นใจในการเปรียบเทียบ และสมมติฐานเบื้องต้นของบุคคลในงานวิจัยนี้ คือ มีความแตกต่างของสีเสื้อผ้าระหว่างบุคคล



รูปที่ 22 การแบ่งบริเวณ โครงร่างบุคคลที่ได้จากการตรวจหา

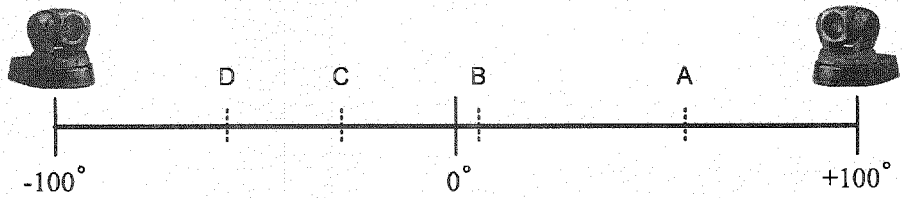
โดยได้แยกบริเวณของโครงร่างบุคคลที่ตรวจหาได้ออกเป็น 3 บริเวณ คือ บริเวณสีเสื้อผ้า บริเวณสีผิว และบริเวณสีผม ดังรูปที่ 22 จะได้ฮิสโทแกรมของทั้งสามส่วนเป็น H_{cloths} H_{skin} และ H_{hair} ดังนั้นเมื่อเปรียบเทียบกับระยะ Bhattacharyya จะได้คะแนนของการเปรียบเทียบซึ่งจะมีการกำหนดค่าขีดเริ่มเปลี่ยนเพื่อใช้ตัดสินบุคคลว่าเป็นบุคคลคนเดิมหรือไม่ ดังสมการที่ 14

$$Bh_{score} < \tau_c \quad (14)$$

โดยที่ Bh_{score} คือ ค่าสัมประสิทธิ์ Bhattacharyya ที่ได้จากการเปรียบเทียบฮิสโทแกรมของแต่ละส่วนบริเวณระหว่างฮิสโทแกรมสีที่ถูกส่งมาจากกล้องวิดีโอตัวที่ 1 กับฮิสโทแกรมสี ณ ขณะนั้นที่รับได้จากกล้องวิดีโอตัวที่ 2 และ τ_c คือ ค่าขีดเริ่มเปลี่ยนในการตัดสินใจสำหรับระบุบุคคล

การพิจารณาทิศทางการเคลื่อนที่ของบุคคลที่สนใจ

อีกลักษณะหนึ่งที่จะถูกนำมาพิจารณาระบุบุคคลในงานวิจัยนี้ คือ ลักษณะทิศทางการเคลื่อนที่ของบุคคลเป้าหมายที่สนใจ เมื่อสมมติฐานการเคลื่อนที่ของบุคคล น่าจะมีเป้าหมายที่แน่นอน เช่น บุคคล ก ต้องการเดินไปยังห้องเรียนตามทางเดินระเบียง แล้วกล้องวิดีโอที่รับภาพทั้งสองก็วางอยู่ในตำแหน่งต้นและปลายของทางเดิน ดังนั้นสมมติว่ากล้องวิดีโอตัวที่ 1 ระบุได้ว่า บุคคลมีทิศทางการเคลื่อนที่ไปทางซ้าย ระบบก็จะส่งข้อมูลทิศทางนี้ไปยังกล้องวิดีโอตัวที่ 2 เพื่อใช้ในการพิจารณาระบุบุคคลเป้าหมาย โดยแบ่งรูปแบบการพิจารณาการเคลื่อนที่ออกเป็น 2 ลักษณะ คือ



รูปที่ 23 การพิจารณาทิศทางการเคลื่อนที่ของบุคคลจากตำแหน่งการเปลี่ยนมุมมองของกล้อง

1. สำหรับกล้องวิดิทัศน์ตัวแรกที่ตรวจพบบุคคลเป้าหมายก่อน บุคคลแรกที่เดินเข้ามาในระบบตรวจหา ไม่ว่าจะเดินเข้ามายังมุมมองกล้องวิดิทัศน์ตัวที่ 1 หรือ ตัวที่ 2 จะถูกกำหนดให้เป็นบุคคลเป้าหมาย และกล้องจะทำการสายและก้มเงยติดตามบุคคลนี้ไปจนสุดขอบเขตที่กล้องวิดิทัศน์จะสามารถเปลี่ยนมุมมองได้ การพิจารณาทิศทางการเคลื่อนที่จะพิจารณาจากแนวทางตำแหน่งการเปลี่ยนมุมมองการรับภาพของกล้องวิดิทัศน์ โดยคำนวณจากการเปลี่ยนตำแหน่งจากตำแหน่งเริ่มต้นของการตรวจหาไปจนถึงตำแหน่งสุดท้าย จากรูปที่ 23 กล้องวิดิทัศน์สามารถเปลี่ยนมุมมองในแนวแกนสายได้ระหว่างมุม -100 องศา ถึง +100 องศา สมมติถ้าตำแหน่งเริ่มต้นของการเปลี่ยนมุมมองกล้องวิดิทัศน์อยู่ที่ตำแหน่ง A และตำแหน่งการเปลี่ยนมุมมองสุดท้ายของกล้องวิดิทัศน์อยู่ที่ D จะพิจารณาว่ามีทิศทางการเคลื่อนที่ไปทางซ้าย (left direction) ของกล้องวิดิทัศน์ ถ้าตำแหน่งเริ่มต้นของการเปลี่ยนมุมมองกล้องวิดิทัศน์อยู่ที่ตำแหน่ง C และตำแหน่งการเปลี่ยน มุมมองสุดท้ายของกล้องวิดิทัศน์อยู่ที่ B จะพิจารณาว่ามีทิศทางการเคลื่อนที่ไปทางขวา (right direction) ของกล้องวิดิทัศน์

2. สำหรับกล้องวิดิทัศน์ตัวที่สองที่ต้องทำการตรวจสอบระบุบุคคล ในกรณีนี้กล้องวิดิทัศน์จะยังไม่เคลื่อนที่จนกว่าจะระบุได้ว่าเป็นบุคคลเป้าหมายที่เคลื่อนที่ผ่านกล้องวิดิทัศน์ตัวแรกมา ดังนั้นจึงใช้วิธีการคำนวณการติดตามเบื้องต้นได้ คือ พิจารณาจากตำแหน่งวัตถุก่อนหน้านี้กับตำแหน่งวัตถุปัจจุบัน ดังนี้

$$\text{if } \begin{cases} X_t^j - X_{t-1}^j > 0 & ; \text{ right direction} \\ X_t^j - X_{t-1}^j < 0 & ; \text{ left direction} \\ X_t^j - X_{t-1}^j = 0 & ; \text{ no move} \end{cases} \quad (15)$$

โดยที่ X_t^{column} คือตำแหน่งใบหน้าของบุคคลเป้าหมาย ณ ปัจจุบัน X_{t-1}^{column} คือตำแหน่งใบหน้าของบุคคลเป้าหมายที่ตรวจหาได้ก่อนหน้านี้ (จะพิจารณาเฉพาะตำแหน่งสดมภ์(column) ของภาพเท่านั้น โดยสดมภ์ที่ 0 จะอยู่ทางซ้ายสุดของภาพ

4. ผลสัมฤทธิ์ของโครงการวิจัย

-พัฒนาและออกแบบระบบการตรวจหาติดตามใบหน้าบุคคลที่สนใจ โดยใช้กล้องวีดิทัศน์ PTZ 2 ตัว ทำงานร่วมกัน ซึ่งระบบสามารถประมวลผลได้ในเวลาจริง ที่อัตราเฟรม 15เฟรมต่อวินาที โดยสามารถตรวจหาและควบคุมให้กล้องวีดิทัศน์ PTZ เปลี่ยนมุมมองเพื่อติดตามบุคคลที่มุ่งเน้นไปที่ใบหน้าบุคคล

-ระบบสามารถระบุบุคคลได้ ในแนวทางขั้นตอนวิธีเบื้องต้นของการสร้างระบบที่สามารถตรวจหา ติดตามและรู้จำบุคคลได้พร้อมกันในเวลาจริง

5. งานที่จะทำต่อไป

-ศึกษาแนวทางใหม่และพัฒนาขั้นตอนวิธีการการระบุบุคคลให้มีประสิทธิภาพมากขึ้น และยังสามารถประยุกต์ใช้กับระบบที่ทำงานในเวลาจริง

-พัฒนาขั้นตอนวิธีการตรวจหาใบหน้าให้มีประสิทธิภาพมากขึ้น จากปัญหาที่เกิดขึ้น ในกรณีที่บุคคลมีการเคลื่อนที่เร็วเกินไป จนส่งผลกระทบต่อผลการตรวจหาการเคลื่อนที่ที่ตรวจหาได้ และพัฒนาให้มีความยืดหยุ่นต่อสิ่งแวดล้อมมากยิ่งขึ้น ในกรณีที่เปลี่ยนสถานที่การประมวลผล

-พัฒนาในส่วนการเขียนโปรแกรม เพื่อลดเวลาในการประมวลผลเมื่อต้องมีการประมวลผลกับสัญญาณวีดิทัศน์จากทั้งสองตัวพร้อมกัน

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

1 จุดประสงค์ของงานวิจัย

- 1.1 พัฒนาวีธีแก้ไขและปกปิดความผิดพลาดเพื่อเพิ่มคุณภาพของสัญญาณวีดิทัศน์ ที่ลดทอนลงจากความผิดพลาดในช่องสัญญาณ เมื่อทำการส่งบน โครงข่ายสื่อสาร ไร้สาย ทั้งในแบบจำลองและสถานการณ์จริง
- 1.2 สร้างระบบการประชุมสัมมนาทางวีดิทัศน์ ในรูปแบบของซอฟต์แวร์เพื่อใช้ในการรับ-ส่งสัญญาณวีดิทัศน์บนช่องสัญญาณ ไร้สาย

2 งานวิจัยที่ได้ทำไปแล้ว

- 2.1 ระบบซอฟต์แวร์จำลองการรับ-ส่ง H.264/MPEG-4 part 10 ร่วมกับตัวเข้ารหัส-ถอดรหัสวีดิทัศน์ มาตรฐาน H.264 AVC/MPEG-4 Part 10 ผ่านช่องสื่อสาร ไร้สายจำลองที่มีความผิดพลาด เพื่อจำลองการเสียหายของข้อมูลในแมโครบล็อก (Macroblock) และศึกษาผลกระทบของความผิดพลาดที่เกิดขึ้นกับคุณภาพของวีดิทัศน์ มีรายละเอียดดังนี้
 - 2.1.1 ใช้การปกปิดความผิดพลาด/การสูญหายของข้อมูลแมโครบล็อก ด้วยข้อมูลจากเฟรมก่อนหน้า
 - 2.1.2 ใช้การทำงานในส่วนการควบคุมอัตรา (Rate-control) ของตัวเข้ารหัส
 - 2.1.3 จำลองผลที่เกิดขึ้นจากการแพร่ (propagate) ของความผิดพลาดภายใน slice
 - 2.1.4 ปรับปรุงการใช้งาน FMO บนช่องสัญญาณ ไร้สายจำลองเพื่อลดผลจากความผิดพลาดในช่องสัญญาณ
 - 2.1.5 ใช้การเข้ารหัส FEC เพื่อลดเพื่อลดผลจากความผิดพลาดในช่องสัญญาณ
- 2.2 ทำการทดลองโดยเปลี่ยนค่าการจัดแผนที่ยกุ่มสไลซ์ด้วยความสามารถการจัดเรียงแมโครบล็อกแบบยืดหยุ่นได้ (FMO) และวัดประสิทธิภาพเชิงปริมาณด้วยอัตราส่วน PSNR
- 2.3 นำเสนอเทคนิคการจัดกลุ่มสไลซ์ด้วยข่าวสารบิตนับ
- 2.4 วัดผลคุณภาพวีดิทัศน์จากการจำลองเพื่อใช้ในการวิเคราะห์ระบบ

3 รายละเอียดโดยย่อของงานวิจัย

งานวิจัยให้ความสำคัญในการแก้ปัญหาหลักของการรับ-ส่งวีดิทัศน์ดิจิทัลบนช่องสัญญาณ ไร้สายอันได้แก่ ความไม่คงที่ของแบนด์วิดท์, การสูญหายของกลุ่มข้อมูล และความผิดพลาดของกลุ่มข้อมูล เพื่อสร้างระบบการลกรหัสวีดิทัศน์ที่มีการปรับตัวได้ตามสภาวะของช่องสัญญาณ และสามารถแก้ไขผลของการสูญหายและความผิดพลาดของกลุ่มข้อมูล ซึ่งจากการศึกษาพบว่าการเข้ารหัสวีดิทัศน์ด้วยมาตรฐาน AVC/MPEG-4 Part 10 สามารถกำหนดโครงสร้างของการจัดเรียงแมโครบล็อกภายในสไลซ์ได้หลายรูปแบบสามารถนำมาประยุกต์ใช้ได้ นอกจากนี้การรับ-ส่งวีดิทัศน์ในบางงานถูกข้อจำกัดเรื่อง

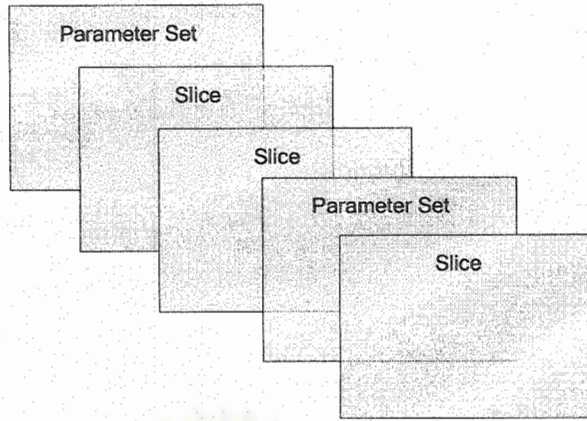
เวลาเข้ามาเกี่ยวข้องกับคือ การรับ-ส่งวิดีโอต้องเป็นแบบเวลาจริง จำนวนรอบของการรับ-ส่งวิดีโอที่ซ้ำกันที่มีการสูญหายของวิดีโอถูกจำกัดให้ทำได้น้อยครั้งมากการใช้งานตัวเข้ารหัส-ถอดรหัส MPEG-4 Part 10 (H.264 AVC)

3.1.1 พื้นฐานตัวเข้ารหัส-ถอดรหัส

มาตรฐาน H.264 (หรืออีกชื่อหนึ่งคือ MPEG-4 part 10) ซึ่งมีประสิทธิภาพการเข้ารหัสสูงกว่ามาตรฐานการบีบอัดวิดีโออื่นก่อนหน้านี้ นั่นคือวิดีโอที่มีอัตราส่วน อัตราบิดต่ออัตราการบิดเบือน (Bit-rate/Distortion ratio) เนื่องจากการเข้ารหัสที่ดีกว่า โครงสร้างข้อมูลของมาตรฐาน H.264 แบ่งเป็นลำดับชั้นต่าง ๆ ได้แก่

- ซับบล็อก คือส่วนพื้นฐานของกระแสข้อมูลวิดีโอ
- แมโครบล็อก (Macroblock) แต่ละแมโครบล็อกประกอบด้วย 1 ซับบล็อกความสว่างขนาด 4x4 จุดภาพ และ 1 ซับบล็อกสีขนาด 2x2 จุดภาพ
- สไลซ์ (Slice) ประกอบด้วย 1 ส่วนหัวสไลซ์ (Slice header) และ 1 ลำดับของแมโครบล็อกที่เรียงต่อกัน และ เฟรม ประกอบจากสไลซ์จำนวนหนึ่งที่รวมกันจนได้ขนาดภาพเต็มภาพ

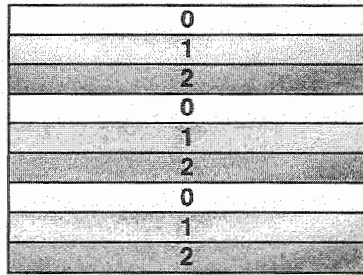
ในส่วนของการเข้ารหัสยังมีการส่งข้อมูล ชุดพารามิเตอร์ (Parameter Set) ประกอบด้วยข้อมูลสำหรับการถอดรหัสซึ่งสามารถใช้สำหรับการถอดรหัส 1 สไลซ์หรือมากกว่า ตัวอย่างของข้อมูลที่เก็บไว้ในชุดพารามิเตอร์เช่น ตัวบ่งชี้วิธีเข้ารหัสเอนโทรปี (Entropy-coding-mode flag) ซึ่งใช้บอกวิธีการที่ใช้เข้ารหัสเอนโทรปีสำหรับซับบล็อกว่าเป็นแบบ CABAC หรือ CAVLC นอกจากนี้ในมาตรฐาน ยังได้กำหนดชั้นของการส่งข้อมูลบนโครงข่าย เรียกว่า หน่วย Network Abstraction Layer (NAL) ประกอบด้วย 1 ส่วนหัว และส่วนที่เหลือซึ่งอาจเป็นได้ทั้ง 1 สไลซ์ หรือ 1 ชุดพารามิเตอร์ NAL ออกแบบมาเพื่อให้เหมาะสมกับทั้งการสื่อสารและการจัดเก็บ ตัวอย่างเช่นไฟล์ H.264 ประกอบด้วย NAL ที่เรียงต่อกัน หากต้องการรับ-ส่งไฟล์ สามารถใช้แต่ละ NAL ในการรับ-ส่งบนโครงข่ายได้ทันที สำหรับการรับ-ส่งบนโครงข่ายที่รับ-ส่งข้อมูลเป็นกลุ่ม (Packet-based network) หรือสามารถแบ่งแต่ละ NAL เป็นส่วนข้อมูล (Payload) ไปใช้ในโครงข่ายที่รับส่งข้อมูลเป็นเฟรม (Frame-based network) ก็ได้



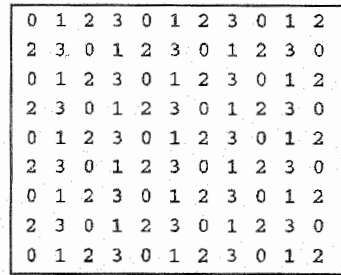
รูปที่ 1 ตัวอย่างของหน่วย NAL

3.1.2 ความสามารถ การจัดแผนที่ยกกลุ่มสไลซ์แบบยืดหยุ่นได้ (Flexible Macroblock Ordering, FMO)

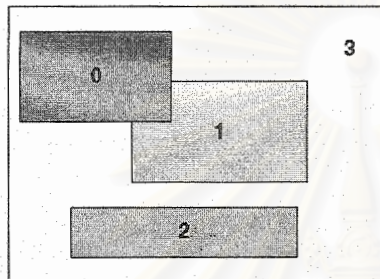
กลุ่มสไลซ์ (Slice group) ประกอบด้วย 1 เซตย่อยของแมโครบล็อกใน 1 เฟรมที่เข้ารหัส และอาจประกอบด้วย 1 สไลซ์หรือมากกว่า สำหรับแต่ละสไลซ์ ที่อยู่ใน 1 กลุ่มสไลซ์ แมโครบล็อกจะถูกเข้ารหัสในลำดับการกวาดแบบแรสเตอร์ (Raster) ถ้าใช้เพียงกลุ่มสไลซ์เดียวสำหรับการเข้ารหัสแต่ละเฟรม จะทำให้แมโครบล็อกทั้งหมดในเฟรมมีการจัดเรียงตามลำดับการกวาดแบบแรสเตอร์ การใช้จำนวนกลุ่มสไลซ์มากกว่า 1 เรียกว่าการจัดเรียงแมโครบล็อกแบบยืดหยุ่น (Flexible Macroblock Ordering) หรือ FMO ช่วยให้สามารถจัดลำดับของการเข้ารหัสแมโครบล็อกสำหรับการถอดรหัสเฟรมได้หลายรูปแบบ ตำแหน่งของแมโครบล็อกหาได้จากแผนที่กลุ่มสไลซ์ (Slice group map) ซึ่งบอกว่าแต่ละแมโครบล็อกอยู่ในกลุ่มสไลซ์ใด แผนที่กลุ่มสไลซ์แต่ละแบบมีลักษณะการจัดกลุ่มแมโครบล็อกและชื่อเรียกตามรูปที่ 6 ตัวอย่างจากรูปที่ 2 แมโครบล็อกจะถูกจัดกลุ่มให้อยู่ในสไลซ์ตามหมายเลขที่อยู่บนแผนที่ ตัวอย่างเช่นรูปที่ 2 (ก) ในเฟรมมีจำนวนสไลซ์ทั้งหมด 3 สไลซ์ได้แก่สไลซ์หมายเลข 0, 1 และ 2 แต่ละสไลซ์มีแมโครบล็อกจำนวน 33 แมโครบล็อก ส่วนรูปที่ 2 (ข) เฟรมมีจำนวนสไลซ์ทั้งหมด 4 สไลซ์ นอกจากแผนที่กลุ่มสไลซ์ทั้ง 6 แบบที่กำหนดไว้ก่อนแล้วยังมีอีกแบบเรียกว่า “Explicit” คือจัดให้แต่ละแมโครบล็อกอยู่ในกลุ่มสไลซ์ใดก็ได้ด้วยการกำหนดเองจากผู้ใช้ที่ตัวเข้ารหัส



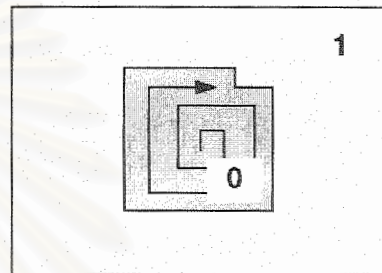
Interleave



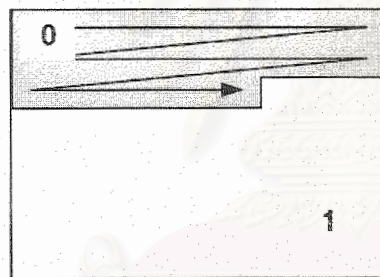
Dispersed



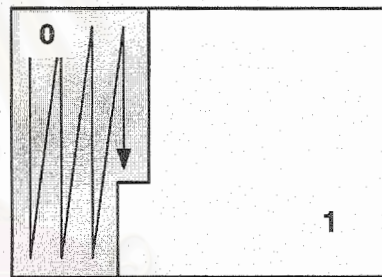
Foreground and background



Box-out



Raster

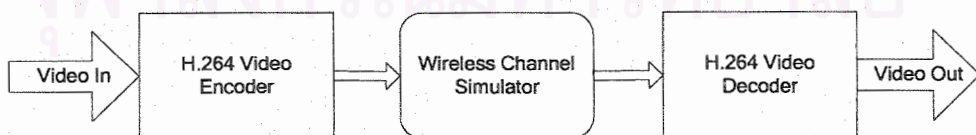


Wipe

รูปที่ 2 การจัดแผนที่กลุ่มสไลซ์แบบต่าง ๆ

3.1.3 โครงข่ายสื่อสารไร้สายและความผิดพลาด

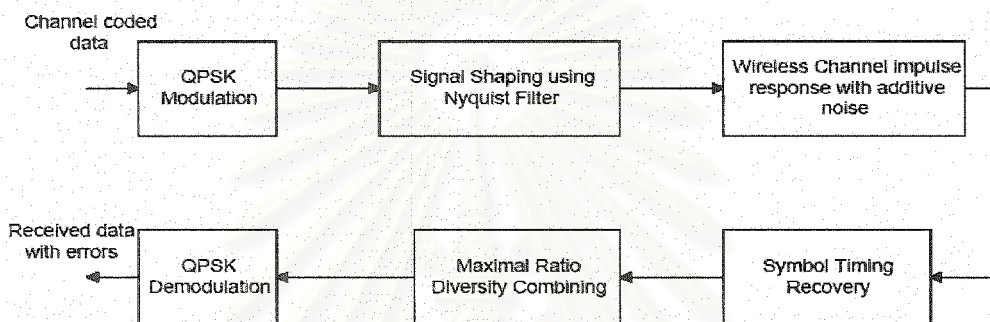
การจำลองการสื่อสารวิทยุบนช่องสัญญาณไร้สายเขียนได้เป็นแผนภาพบล็อกดังรูปที่ 3



รูปที่ 3 แผนภาพบล็อกของการจำลองการสื่อสาร

ในที่นี้เลือกใช้การจำลองช่องสัญญาณไร้สายชนิด PACS (Personal Access Communication Services) [13] ที่สามารถให้วิสัยสามารถ (Throughput) ได้ประมาณ 32 กิโลบิตต่อวินาที ซึ่งเหมาะสมกับการใช้งานสื่อสารวิทยุ ทำการจำลองด้วยค่าพารามิเตอร์ทั่วไป ระบบ PACS

เป็นมาตรฐานของ North American และเป็นช่องสัญญาณไร้สายชนิดหนึ่งจากหลายชนิดที่มีใช้งานได้ในอัตราบิตระดับนี้ เช่น DECT (Digital European Cordless Telephone) และ PHS (Personal Handy phone System) ความยาวของเฟรมพาหะ PACS (Frame carries) เท่ากับ 120 บิต ประกอบด้วย 80 บิตข้อมูลสำหรับผู้ใช้ และ 40 บิตสำหรับส่วนหัว (และบิต CRC) . ค่าประวิงเวลา “Round trip” ของระบบ PACS มีค่าน้อย อยู่ในระดับ 10 มิลลิวินาที ซึ่งเหมาะสมกับงานสื่อสารวิทยุทัศน์สองทางที่ต้องการการประวิงเวลาดำ ตัวจำลองช่องสัญญาณที่ใช้เป็นการจำลองช่องสัญญาณที่เรียกว่า “Rayleigh fading channel” มีขั้นตอนการทำงานของตัวจำลองช่องสัญญาณแสดงได้เป็นแผนภาพดังรูปที่ 4



รูปที่ 4 แผนภาพการจำลองช่องสัญญาณ

3.1.4 การเข้ารหัส Forward Error Correction (FEC)

ความผิดพลาดที่เกิดจากการสื่อสารสามารถตรวจสอบและแก้ไขได้ด้วยการเข้ารหัส FEC ซึ่งใช้การเพิ่มข้อมูลซ้ำซ้อนให้ได้คำรหัส (Code word) ที่ความน่าจะเป็นของการถอดรหัสได้เป็นคำรหัสอื่นลดลง การเข้ารหัส FEC แบ่งได้เป็น 2 ชนิดมีชื่อเรียกต่างกัน ได้แก่ “Linear block code” และ “Convolution code” ที่มีข้อแตกต่างกันในรายละเอียด แต่ตั้งอยู่บนพื้นฐานเดียวกัน ในการเข้ารหัส Linear block code ข้อมูลที่ต้องการส่งจะถูกแบ่งส่วนย่อยเป็นบล็อกของ k บิต สำหรับการเข้ารหัส (n, k) บิตส่วนซ้ำซ้อนจำนวน $r = n - k$ จะถูกเพิ่มเข้าไปกับ k ทำให้ได้จำนวนบิตทั้งหมด n บิต แต่ละบล็อกจะมีการคำนวณความซ้ำซ้อน หรือที่เรียกว่า “Parity symbol” และได้อัตราส่วนการเข้ารหัส $cr = k/n$ ความสามารถในการแก้ความผิดพลาดของการเข้ารหัส (n, k) ขึ้นกับค่าระยะแฮมมิงน้อยสุด (Minimum hamming distance, d_{min}) ซึ่งควรมีค่ามากที่สุด ระยะแฮมมิงของคำรหัสไบนารี (Binary codeword) คือจำนวนของบิตที่ต่างกันระหว่างคำรหัส สำหรับรหัสที่มีระยะแฮมมิง d_{min} จำนวนบิตผิดพลาดที่สามารถแก้ไขได้จะเท่ากับอย่างน้อย น้อย $t = \lfloor (d_{min} - 1) / 2 \rfloor$ บิต เมื่อสัญลักษณ์ $\lfloor x \rfloor$ หมายถึงหมายถึงจำนวนเต็มมากที่สุดที่ไม่เกิน x

3.2 ตัวอย่างผลการจำลอง

3.2.1 ผลการจำลองเพื่อศึกษาผลกระทบของความผิดพลาดในช่องสื่อสารที่มีต่อวิคิตส์ การจำลองกลุ่มนี้ให้ความสนใจเฉพาะผลจากความผิดพลาดเพียงอย่างเดียว การเข้ารหัสไม่มีการใช้งาน FMO และ FEC ร่วมด้วย และใช้การควบคุมอัตราที่ 32 กิโลบิตต่อวินาทีซึ่งเท่ากับพิกัดแบนด์วิคิตส์ที่กำหนดไว้ของช่องสื่อสาร

กรณีเฟดคิงช้า 1Hz				
ชื่อลำดับภาพ	จำนวนแมโครบล็อกที่ถอดรหัสไม่ได้	ค่าเฉลี่ย PSNR (dB)		
		ไม่มีความผิดพลาด	มีความผิดพลาด	ผลต่าง
akiyo	1547	42.28	30.71	11.57
carphone	1380	33.57	23.04	10.53
claire	1346	42.78	30.82	11.97
foreman	1499	32.50	17.11	15.39

กรณีเฟดคิงเร็ว 40Hz				
ชื่อลำดับภาพ	จำนวนแมโครบล็อกที่ถอดรหัสไม่ได้	ค่าเฉลี่ย PSNR (dB)		
		ไม่มีความผิดพลาด	มีความผิดพลาด	ผลต่าง
akiyo	5367	42.28	27.63	14.65
carphone	5634	33.57	19.01	14.56
claire	5269	42.78	24.31	18.47
foreman	5740	32.50	15.12	17.39

3.2.2 ผลการจำลองเพื่อศึกษาผลกระทบการใช้งาน FMO ที่มีผลต่อส่วนการทำนายการเคลื่อนที่ภายในตัวเข้ารหัส โดยไม่มีความผิดพลาดในช่องสื่อสารและมีการควบคุมอัตราที่ 32 กิโลบิตต่อวินาทีซึ่งเท่ากับพิกัดแบนด์วิคิตส์ที่กำหนดไว้ของช่องสื่อสาร กลุ่มของการจำลองแบ่งตามรูปแบบของแผนที่กลุ่มสไลซ์ที่ใช้ ทำการจำลองที่จำนวนกลุ่มสไลซ์ต่อภาพต่างกัน

3.2.2.1 การเข้ารหัสด้วยแผนที่กลุ่มสไลซ์แบบคงที่ใช้รูปแบบแผนที่กลุ่มสไลซ์แบบเดียวกันสำหรับการเข้ารหัสทุกเฟรม

ชื่อแผนทีกลุ่มสไลซ์ / จำนวนกลุ่มสไลซ์ต่อภาพ

0 0 0 0 0 0 0 0 0 0 0 0			
0 0 0 0 0 0 0 0 0 0 0 0			
0 0 0 0 0 0 0 0 0 0 0 0			
0 0 0 0 0 0 0 0 0 0 0 0			
0 0 0 0 0 0 0 0 0 0 0 0			
0 0 0 0 0 0 0 0 0 0 0 0			
0 0 0 0 0 0 0 0 0 0 0 0			
0 0 0 0 0 0 0 0 0 0 0 0			
0 0 0 0 0 0 0 0 0 0 0 0			
0 0 0 0 0 0 0 0 0 0 0 0			
0 0 0 0 0 0 0 0 0 0 0 0			

FMO

0 0 0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0 0 0	0 0 0 1 1 1 1 1 1 1 1 1
0 0 0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0 0 0	1 1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 1 1 1
0 0 0 0 0 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 2 2 2 2
1 1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 1 1 1	2 2 2 2 2 2 2 2 2 2 2 2
1 1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 1 1 1	2 2 2 2 2 2 2 2 2 2 2 2	2 2 2 2 2 2 2 2 2 2 2 2
1 1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 1 1 1	2 2 2 2 2 2 2 2 2 2 2 2	3 3 3 3 3 3 3 3 3 3 3 3
1 1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 1 1 1	2 2 2 2 2 2 2 2 2 2 2 2	3 3 3 3 3 3 3 3 3 3 3 3

	Interleaved/2	Interleaved/3	Interleaved/4
0 0 0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 1 1	0 0 0 0 0 0 0 0 1 1 1 1	0 0 0 1 1 1 1 1 1 1 1 1	0 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 2 2 2 2 2 2	1 1 2 2 2 2 2 2 2 2 2 2
1 1 1 1 1 1 1 2 2 2 2 2	1 2 2 2 2 2 2 2 2 2 2 2	2 2 2 2 2 2 2 2 2 2 3 3	2 2 2 2 3 3 3 3 3 3 3 3
2 2 2 2 2 2 2 2 2 2 2 2	2 2 2 2 2 2 2 2 2 2 2 2	3 3 3 3 3 3 3 3 3 3 3 3	3 3 3 3 4 4 4 4 4 4 4 4
2 2 2 2 3 3 3 3 3 3 3 3	3 3 3 3 3 3 3 3 3 3 3 3	4 4 4 4 4 4 4 4 4 4 4 4	4 4 4 4 4 5 5 5 5 5 5 5 5
3 3 3 3 3 3 3 3 3 3 3 3	3 3 4 4 4 4 4 4 4 4 4 4	4 4 4 4 5 5 5 5 5 5 5 5	5 5 5 5 5 5 5 5 6 6 6 6
3 3 4 4 4 4 4 4 4 4 4 4	4 4 4 4 4 4 4 4 4 5 5 5	5 5 5 5 5 5 5 5 6 6 6 6	6 6 6 6 6 6 6 6 6 6 7 7
4 4 4 4 4 4 4 4 4 4 4 4	5 5 5 5 5 5 5 5 5 5 5 5	6 6 6 6 6 6 6 6 6 6 6 6	7 7 7 7 7 7 7 7 7 7 7 7

Interleaved/5

	Interleaved/6	Interleaved/7	Interleaved/8
0 1 0 1 0 1 0 1 0 1 0 1	0 1 0 1 0 1 0 1 0 1 0 1	0 1 2 0 1 2 0 1 2 0 1	0 1 2 3 0 1 2 3
1 0 1 0 1 0 1 0 1 0 1 0	1 0 1 0 1 0 1 0 1 0 1 0	2 0 1 2 0 1 2 0 1 2 0	3 0 1 2 3 0 1 2
0 1 0 1 0 1 0 1 0 1 0 1	0 1 0 1 0 1 0 1 0 1 0 1	1 2 0 1 2 0 1 2 0 1 2	2 3 0 1 2 3 0 1
1 0 1 0 1 0 1 0 1 0 1 0	1 0 1 0 1 0 1 0 1 0 1 0	0 1 2 0 1 2 0 1 2 0 1	1 2 3 0 1 2 3 0
0 1 0 1 0 1 0 1 0 1 0 1	0 1 0 1 0 1 0 1 0 1 0 1	2 0 1 2 0 1 2 0 1 2 0	0 1 2 3 0 1 2 3
1 0 1 0 1 0 1 0 1 0 1 0	1 0 1 0 1 0 1 0 1 0 1 0	1 2 0 1 2 0 1 2 0 1 2	3 0 1 2 3 0 1 2
0 1 0 1 0 1 0 1 0 1 0 1	0 1 0 1 0 1 0 1 0 1 0 1	0 1 2 0 1 2 0 1 2 0 1	2 3 0 1 2 3 0 1
1 0 1 0 1 0 1 0 1 0 1 0	1 0 1 0 1 0 1 0 1 0 1 0	2 0 1 2 0 1 2 0 1 2 0	1 2 3 0 1 2 3 0
0 1 0 1 0 1 0 1 0 1 0 1	0 1 0 1 0 1 0 1 0 1 0 1	1 2 0 1 2 0 1 2 0 1 2	0 1 2 3 0 1 2 3

Dispersed /2

Dispersed /3

Dispersed /4

<pre> 1 2 3 4 0 1 2 3 4 0 2 3 4 0 1 2 3 4 0 1 3 4 0 1 2 3 4 0 1 2 4 0 1 2 3 4 0 1 2 3 0 1 2 3 4 0 1 2 3 4 1 2 3 4 0 1 2 3 4 0 2 3 4 0 1 2 3 4 0 1 3 4 0 1 2 3 4 0 1 2 4 0 1 2 3 4 0 1 2 3 </pre>	<pre> 0 1 2 3 4 5 0 1 2 3 4 5 0 1 2 3 4 5 0 1 2 3 4 5 0 1 2 3 4 5 0 1 2 3 4 5 0 1 2 3 4 5 0 1 2 3 4 5 0 1 2 3 4 5 0 1 2 3 4 5 0 1 2 3 4 5 0 1 2 3 4 5 0 1 2 3 4 5 0 1 2 3 4 5 0 1 2 3 4 5 0 1 2 3 4 5 0 1 2 </pre>	<pre> 0 1 2 3 4 5 6 0 1 2 3 4 5 6 0 1 2 3 4 5 6 0 1 2 3 4 5 6 0 1 2 3 4 5 6 0 1 2 3 4 5 6 0 1 2 3 4 5 6 0 1 2 3 4 5 6 0 1 2 3 4 5 6 0 1 2 3 4 5 6 0 1 2 3 4 5 6 0 1 2 3 4 5 6 0 1 2 3 4 5 6 0 1 2 3 4 5 6 0 </pre>	<pre> 0 1 2 3 4 5 6 7 0 3 4 5 6 7 0 1 2 3 6 7 0 1 2 3 4 5 6 1 2 3 4 5 6 7 0 1 4 5 6 7 0 1 2 3 4 7 0 1 2 3 4 5 6 7 2 3 4 5 6 7 0 1 2 5 6 7 0 1 2 3 4 5 0 1 2 3 4 5 6 7 0 </pre>
--	--	--	--

dispersed /5 Dispersed /6 Dispersed /7 Dispersed /8

<pre> 0 0 0 0 0 0 1 1 1 1 1 0 0 0 0 0 0 1 1 1 1 1 0 0 0 0 0 0 1 1 1 1 1 0 0 0 0 0 0 1 1 1 1 1 0 0 0 0 0 1 1 1 1 1 1 0 0 0 0 0 1 1 1 1 1 1 0 0 0 0 0 1 1 1 1 1 1 0 0 0 0 0 1 1 1 1 1 1 0 0 0 0 0 1 1 1 1 1 1 0 0 0 0 0 1 1 1 1 1 1 </pre>	<pre> 0 0 0 1 1 1 1 2 2 2 2 0 0 0 1 1 1 1 2 2 2 2 0 0 0 1 1 1 1 2 2 2 2 0 0 0 0 1 1 1 2 2 2 2 0 0 0 0 1 1 1 2 2 2 2 0 0 0 0 1 1 1 2 2 2 2 0 0 0 0 1 1 1 2 2 2 2 0 0 0 0 1 1 1 1 2 2 2 0 0 0 0 1 1 1 1 2 2 2 0 0 0 0 1 1 1 1 2 2 2 </pre>	<pre> 0 0 0 1 1 1 2 2 2 2 0 0 0 1 1 1 2 2 2 2 0 0 0 1 1 1 2 2 2 2 0 0 0 1 1 1 2 2 2 3 0 0 0 1 1 1 2 2 3 0 0 0 1 1 1 2 2 3 0 0 0 1 1 1 2 2 3 0 0 1 1 1 1 2 2 3 0 0 1 1 1 1 2 2 3 </pre>
--	--	--

Wipe/2 Wipe/3 Wipe/4

<pre> 0 0 1 1 2 2 3 3 4 4 0 0 1 1 2 2 3 3 4 4 0 1 1 1 2 2 3 3 4 4 0 1 1 1 2 2 3 3 4 4 0 1 1 2 2 2 3 3 4 4 0 1 1 2 2 2 3 3 4 4 0 1 1 2 2 3 3 3 4 4 0 1 1 2 2 3 3 3 4 4 0 1 1 2 2 3 3 3 4 4 </pre>	<pre> 0 1 1 2 2 3 3 4 4 5 5 0 0 1 2 2 3 3 4 4 5 5 0 0 1 1 2 2 3 3 4 4 5 5 0 0 1 1 2 2 3 3 4 4 5 5 0 0 1 1 2 2 3 3 4 4 5 5 0 0 1 1 2 2 3 3 4 4 5 0 0 1 1 2 2 3 3 4 4 5 0 0 1 1 2 2 3 3 4 4 5 0 0 1 1 2 2 3 3 4 4 5 </pre>	<pre> 0 1 1 2 2 3 3 5 5 6 6 0 1 1 2 2 3 3 5 5 6 6 0 1 1 2 2 3 3 4 4 5 6 6 0 1 1 2 2 3 3 4 4 5 6 6 0 0 1 2 2 3 3 4 4 5 6 6 0 0 1 2 2 3 3 4 4 5 6 6 0 0 1 2 3 3 4 4 5 5 6 0 0 1 2 3 3 4 4 5 5 6 0 0 1 1 3 3 4 4 5 5 6 </pre>
--	--	--

Wipe/5 Wipe/6 Wipe/7 Wipe/8

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

กรณีเฟดดิ้งช้า 1Hz									
ชื่อลำดับภาพ	ชื่อแผนที่กลุ่ม สไลซ์	ค่าเฉลี่ย PSNR							
		จำนวนกลุ่มสไลซ์ต่อภาพ							
		1	2	3	4	5	6	7	8
akiyo	Interleaved	42.28	42.14	41.98	41.88	41.71	41.59	41.46	41.38
	Dispersed	42.28	41.99	41.87	41.76	41.63	41.54	41.48	41.33
	Wipe	42.28	42.08	41.94	41.95	41.69	41.48	41.54	41.17
carphone	Interleaved	33.57	33.39	33.23	33.04	32.87	32.61	32.44	32.24
	Dispersed	33.57	33.01	32.89	32.64	32.49	32.34	32.19	32.04
	Wipe	33.57	33.45	33.30	33.14	33.02	32.80	32.67	32.59
clair	Interleaved	42.78	42.62	42.44	42.30	42.13	42.03	41.80	41.59
	Dispersed	42.78	42.33	42.17	42.03	41.93	41.80	41.66	41.53
	Wipe	42.78	42.66	42.48	42.25	42.17	42.11	41.98	41.89
foreman	Interleaved	32.50	32.25	32.06	31.84	31.56	31.35	31.14	30.93
	Dispersed	32.50	31.62	31.37	30.99	30.79	30.75	30.59	30.40
	Wipe	32.50	32.35	31.73	31.83	31.62	31.44	31.32	31.19

3.2.3 การจำลองเพื่อศึกษาผลจากการใช้งาน FMO เพื่อลดผลของความผิดพลาดในช่องสื่อสารที่มีต่อวีดิทัศน์ โดยการจำลองกลุ่มนี้ให้ความสนใจเฉพาะผลจากความผิดพลาดและมีการใช้งาน FMO ร่วมด้วยใช้การควบคุมอัตราที่ 32 กิโลบิตต่อวินาทีซึ่งเท่ากับพิกัดแบนด์วิดท์ที่กำหนดไว้ของช่องสื่อสาร ผลการจำลองแบ่งเป็น 2 กรณีได้แก่ กรณีเฟดดิ้งช้า และกรณีเฟดดิ้งเร็ว การจำลองนี้สนใจให้ความสำคัญกับรูปแบบและจำนวนกลุ่มสไลซ์ต่อภาพว่ามีความสัมพันธ์อย่างไรกับคุณภาพของวีดิทัศน์

กรณีเฟดคิงช้า 1Hz										
ชื่อลำดับ ภาพ	ชื่อแผนที่ กลุ่มสไลซ์	ค่าเฉลี่ย PSNR								
		ไม่มี ความ ผิดพลาด	จำนวนกลุ่มสไลซ์ต่อภาพ							
			1	2	3	4	5	6	7	8
akiyo	Interleaved	42.28	30.71	30.67	30.54	31.99	31.23	31.50	30.67	30.40
	Dispersed		30.71	28.90	31.51	32.05	33.34	32.99	33.09	32.18
	Wipe		30.71	31.35	30.66	31.61	31.46	33.70	33.61	30.65
carphone	Interleaved	33.57	23.04	22.61	23.71	22.86	23.49	24.02	23.19	23.91
	Dispersed		23.04	22.82	23.26	23.47	23.97	23.70	23.82	23.81
	Wipe		23.04	22.45	22.61	23.59	23.68	22.73	23.23	23.08
claire	Interleaved	42.78	30.82	29.26	29.14	28.03	29.25	30.37	31.60	31.49
	Dispersed		30.82	29.70	30.72	31.47	30.51	29.81	30.61	31.49
	Wipe		30.82	30.15	30.37	29.63	30.16	30.55	29.19	29.01
foreman	Interleaved	32.50	17.11	17.97	18.23	17.73	18.28	18.77	19.11	19.06
	Dispersed		17.11	17.81	18.08	18.14	18.07	18.24	18.96	18.25
	Wipe		17.11	18.45	17.76	18.33	18.20	18.35	17.97	18.30

กรณีเฟดคิงช้า 1Hz										
ชื่อลำดับ ภาพ	ชื่อแผนที่กลุ่ม สไลซ์	จำนวนแมโครบล็อกที่ถอดรหัสไม่ได้								
		ไม่มี ความ ผิดพลาด	จำนวนกลุ่มสไลซ์ต่อภาพ							
			1	2	3	4	5	6	7	8
akiyo	Interleaved	9900	1547	1256	908	916	829	692	708	812
	Dispersed		1547	1171	1136	967	941	915	792	836
	Wipe		1547	1253	794	655	601	600	524	474

carphone	Interleaved	9900	1380	1286	1109	1143	1012	923	794	888
	Dispersed		1380	1199	1012	955	909	903	779	790
	Wipe		1380	1159	1137	945	879	834	775	744
claire	Interleaved	9900	1346	1275	1042	929	797	723	769	773
	Dispersed		1346	1151	1143	942	903	864	795	724
	Wipe		1346	1149	729	745	558	530	579	471
foreman	Interleaved	9900	1499	1287	1082	914	984	878	844	836
	Dispersed		1499	1066	1074	905	935	790	783	730
	Wipe		1499	1190	1049	930	964	927	825	837

3.2.4 ผลการจำลองนำเสนอเทคนิคการจัดกลุ่มสไลซ์ด้วยข่าวสารบีตนับ

จากการทดลองใช้งาน FMO ได้จำนวนแมโครบล็อกที่ไม่สามารถถอดรหัสได้ลดลง ทั้งในกรณีที่ใช้แผนที่ยกุ่มสไลซ์มาตรฐานและเมื่อใช้กับวิธีการจัดแผนที่ยกุ่มสไลซ์ด้วยข้อมูลจำนวนบิต (explicit) และค่าคุณภาพ PSNR จากบทความ (2.4.1)

TABLE I. AMOUNT OF UNDECODABLE MACROBLOCK IN FRAME FOR EACH MAP TYPE USED (8 SLICES GROUP) SLOW FADING.

Sequence	Map			
	One slice	Interleave	Dispersed	Proposed
akiyo	1547	812	836	720
bridge_clo se	1028	458	462	387
carphone	1380	888	790	615
foreman	1235	858	719	668

TABLE II. AMOUNT OF UNDECODABLE MACROBLOCK IN FRAME FOR EACH MAP TYPE USED (8 SLICES GROUP) FAST FADING.

Sequence	Map			
	One slice	Interleave	Dispersed	Proposed
akiyo	5367	1535	1355	1205
bridge_clos e	4398	972	789	710
carphone	5634	1390	1233	1113
foreman	5903	1337	1232	1193

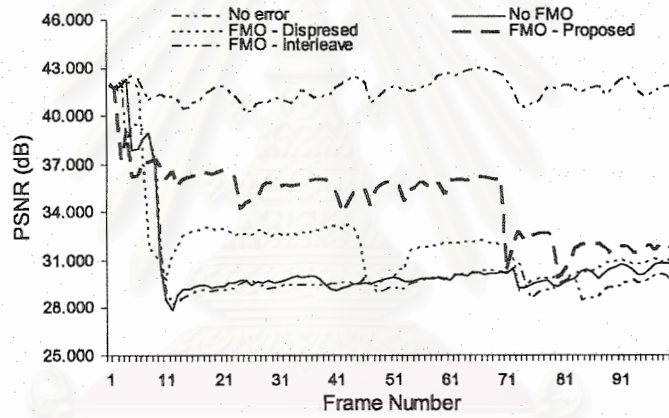


Figure 4. PSNR comparison of Akiyo sequence for slow-fading channel.

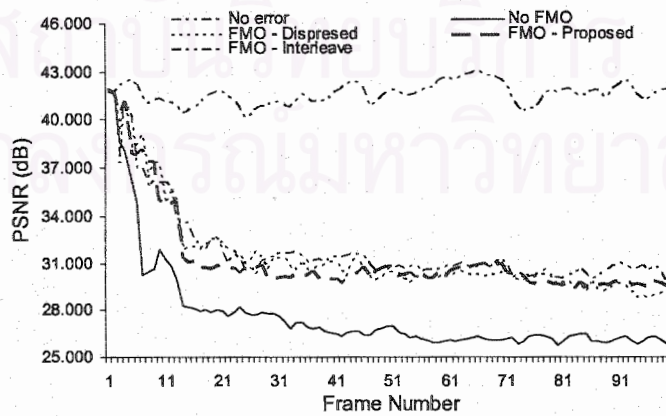


Figure 5. PSNR comparison of Akiyo sequence for fast-fading channel.

4 งานที่จะทำต่อไป

- 4.1 ใช้การปรับค่า FMO และ FEC ด้วยการเพิ่มแบบจำลองช่องสัญญาณ ไร้สายในตัวเข้ารหัส
- 4.2 ใช้การปรับค่าและร่วมกับการใช้งานความสามารถอื่นของตัวเข้ารหัส/ถอดรหัสเพิ่มคุณภาพของวิดีโอ



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ค

1. จุดประสงค์ของงานวิจัย

1.1 ศึกษาคุณลักษณะเฉพาะของความสัมพันธ์ระหว่าง อัตราบิต (Bit rate) กับ ค่าความเพี้ยน (Disatortion) ของการเข้ารหัส โดยอาศัยความรู้พื้นฐานจากคุณสมบัติทางสถิติของสัญญาณวิดีโอต้นฉบับ (Video source) เพื่อใช้หาแบบจำลองทางคณิตศาสตร์ระหว่างอัตราบิตและค่าความเพี้ยน -Rate)

(D Model-R :Distortion model ที่มีความเหมาะสมมากที่สุด เพื่อให้การเข้ารหัสสัญญาณวิดีโอต้นฉบับตามมาตรฐาน AVC/264.H มีประสิทธิภาพสูงสุด

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

1.3 ศึกษาและพัฒนาทฤษฎีสำหรับกรรมวิธีการควบคุมอัตรา โดยนำเอาสถานะของช่องสัญญาณไร้สายมาพิจารณาร่วมกับแบบจำลอง D-R และค่าความเพี้ยนที่เกิดขึ้นในระหว่างการส่งผ่านช่องสัญญาณไร้สาย

2. งานวิจัยที่ได้ทำไปแล้ว

2.1 ศึกษาคุณลักษณะเฉพาะของความสัมพันธ์ระหว่างอัตราบิต (Bit rate) กับค่าระดับการควอนไทซ์ (Quantization step size) ในการเข้ารหัสสัญญาณวิดีโอต้นฉบับตามมาตรฐาน AVC/264.H โดยพิจารณาจากลักษณะการแจกแจงความน่าจะเป็น (Probability distribution) ของสัมประสิทธิ์ DCT ในแต่ละลำดับภาพวิดีโอต้นฉบับ ดังนี้

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

-เปรียบเทียบความสัมพันธ์ระหว่างอัตราบิต กับค่าระดับการควอนไทซ์ หลังการเข้ารหัสจริงตามมาตรฐาน 264.H ของภาพชนิด Intra และ Inter frame เปรียบเทียบกับแบบจำลองของอัตราบิตในรูปแบบฟังก์ชันของระดับการควอนไทซ์ ในรูปแบบของการแจกแจงความน่าจะเป็นแบบโคชี แล้วพบว่ามีความใกล้เคียงกันมาก

2.2 ทำการปรับปรุงการเข้ารหัสสัญญาณวิดีโอต้นฉบับในส่วนของการควบคุมอัตรา ในมาตรฐาน AVC/264.H โดยใช้ Model-Cauchy Rate โดยพิจารณาว่าการแจกแจงความน่าจะเป็นของการแปลงสัมประสิทธิ์ DCT มีลักษณะเป็นการแจกแจงแบบ Cauchy มาใช้ในการคำนวณหาค่า Quantization step size แทนการใช้ Quadratic Model ในมาตรฐาน AVC/264.H ซึ่งเป็น โมเดลที่ได้จากการตั้งสมมุติฐานว่าการแจกแจงความน่าจะเป็นของการแปลงสัมประสิทธิ์ DCT มีลักษณะการแจกแจงเป็นแบบ Laplacian ซึ่ง

จากผลการทดลองพบว่าการใช้ Cauchy Rate- Model สามารถปรับปรุงค่า PSNRของสัญญาณวิดีโอที่สนได้เพิ่มขึ้น

2.3 วิเคราะห์หาสมการทางคณิตศาสตร์ จาก Langrange Optimization Technique โดยอาศัยความสัมพันธ์ระหว่าง Model-Cauchy Rate และ Model-Cauchy Distortion ในการหาสมการที่เหมาะสมที่สุดสำหรับการคำนวณหาค่า Quantization step size เพื่อให้ได้จำนวนบิตตามอัตราบิตเป้าหมายที่กำหนด โดยที่มีความเพี้ยนเกิดขึ้นน้อยที่สุด

2.4 นำ Distortion Model-Rate ที่ได้จากข้อ 2.2 มาประยุกต์ใช้ในกรณีวิธีควบคุมอัตราตามมาตรฐาน AVC/264.H ในระดับหน่วยพื้นฐานย่อย (Basic unit) ซึ่งจากผลการทดลองพบว่าการใช้ Cauchy Rate- Distortion Model สามารถปรับปรุงค่า PSNRของสัญญาณวิดีโอที่สนได้เพิ่มขึ้น โดยที่ความแปรปรวนของค่า PSNR น้อยลง

3. รายละเอียดของงานวิจัย

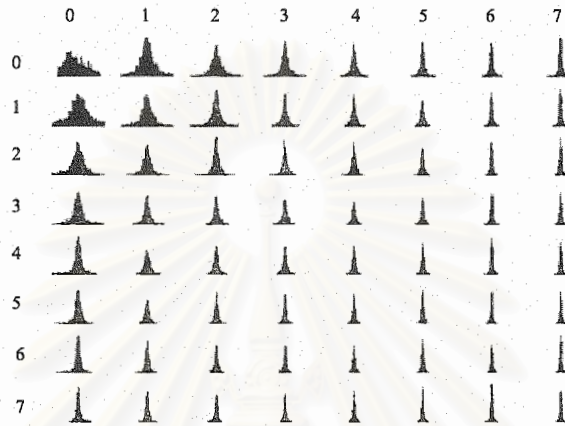
งานวิจัยให้ความสำคัญในการแก้ปัญหาหลักของการรับ-ส่งวิดีโอที่สนดิจิทัลบนช่องสัญญาณ ไร้สายอันได้แก่ ความไม่คงที่ของแบนด์วิดท์, การสูญหายของกลุ่มข้อมูล และความเพี้ยนของข้อมูลในระหว่างการส่ง เพื่อสร้างระบบการเข้ารหัสวิดีโอที่สนที่มีการปรับตัวได้ตามสถานะของช่องสัญญาณ โดยมุ่งเน้นการพัฒนาไปที่กรรมวิธีการควบคุมอัตราสำหรับการเข้ารหัสสัญญาณวิดีโอที่สนตามมาตรฐาน AVC/264.H ซึ่งจะทำการหาแบบจำลองทางคณิตศาสตร์ที่เหมาะสม โดยใช้ Langrange Optimization Technique เพื่อใช้แสดงความสัมพันธ์ระหว่างอัตราบิต และค่าความเพี้ยนที่เกิดจากการเข้ารหัส แล้วทำการปรับปรุงกลไกในส่วนของกรรมวิธีการควบคุมอัตราใหม่ให้มีความสอดคล้องกับ , Distortion Model-Rate การพิจารณาในส่วนของ Low โดยนำเอาข้อมูลจากสถานะบัฟเฟอร์ของตัวเข้ารหัสมาใช้ในการพิจารณาคัดเลือก จำนวนบิตเป้าหมายในกรรมวิธีการควบคุมอัตรา ซึ่งรายละเอียดของงานวิจัยสามารถสรุปได้ดังนี้

การวิเคราะห์ฟังก์ชันความหนาแน่นความน่าจะเป็น (PDF :Probability density function)ของภาพ

ศึกษาคุณลักษณะเฉพาะของความสัมพันธ์ระหว่างอัตราบิต (Bit rate) กับค่าระดับการควอนไทซ์ (Quantization step size) ในการเข้ารหัสสัญญาณวิดีโอที่สนตามมาตรฐาน AVC/264.H โดยพิจารณาจากลักษณะการแจกแจงความน่าจะเป็น (Probability distribution) ของสัมประสิทธิ์ DCT ในแต่ละลำดับภาพวิดีโอที่สน ซึ่งมีรายละเอียดดังนี้

1. วิเคราะห์ฟังก์ชันความหนาแน่นความน่าจะเป็น (PDF :Probability density function)ของภาพ
วิเคราะห์ฟังก์ชันความหนาแน่นความน่าจะเป็น (PDF) ของ สัมประสิทธิ์ชนิด -AC) AC (Coefficients ของภาพหลังการแปลงสัมประสิทธิ์ DCT โดยใช้ลำดับภาพต่างๆ กันดังนี้ Akiyo sequence (pixels 176x 144 :QCIF) Carphone sequence ,(pixels 176x 144 :QCIF) และ :CIF) Tempeete sequence (pixels 352x 288ทำการทดลองโดยแบ่งออกเป็น 2ประเภท ตามชนิดของเฟรม คือ Intra frame และ Nonintra frame

1. ทำการแปลงสัมประสิทธิ์ DCT โดยแบ่งภาพออกเป็นบล็อกขนาด pixels 8x 8 แล้วใช้โปรแกรม Matlab เก็บค่าทางสถิติของสัมประสิทธิ์ DCT ของ 100 เฟรมแรกของแต่ละลำดับภาพ แล้วทำการพล็อตกราฟเพื่อดู PDF ของสัมประสิทธิ์ DCT ที่ตำแหน่งบล็อกเดียวกัน ดังตัวอย่างในรูปที่ 1



รูปที่ 1 แสดง PDF ของแต่ละสัมประสิทธิ์ DCT ของ block ขนาด 8x8

2. จากกราฟแสดงลักษณะการกระจายความหนาแน่นความน่าจะเป็นของ สัมประสิทธิ์ DCT เฉพาะ -AC

Coefficients ทำการ fitting curve โดยแบ่งการพิจารณาออกเป็น 2 กรณีคือ

- พิจารณาจาก Laplacian Distribution จากสมการ

$$P(X) = \frac{1}{2b} \exp\left(-\frac{|x - \mu|}{b}\right)$$

μ = ค่า mean ของข้อมูล

- พิจารณาจาก Cauchy Distribution จากสมการ

$$P(X) = \frac{1}{\pi} \frac{\lambda}{\lambda^2 + (x - x_0)^2}$$

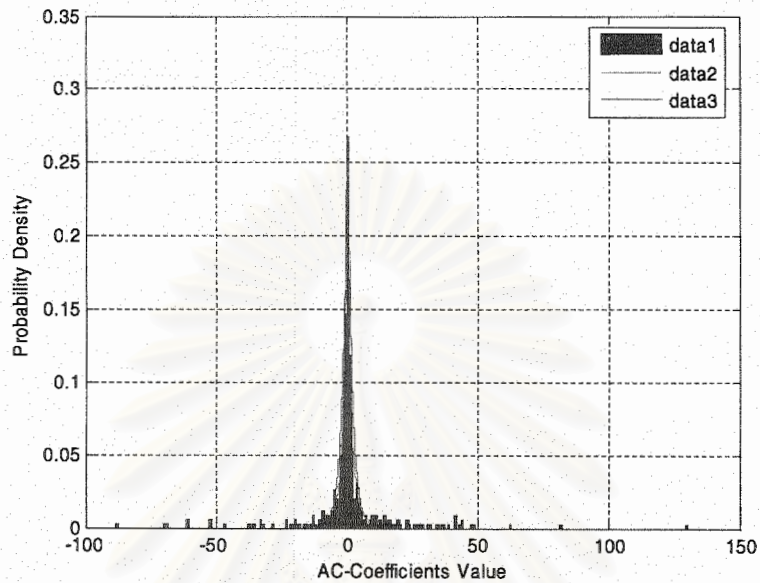
โดยทำการปรับค่าพารามิเตอร์ b สำหรับ Laplacian Distribution และ พารามิเตอร์ λ สำหรับ

Cauchy

Distribution

กรณี Intra frame สามารถแสดงลักษณะการกระจายความหนาแน่นความน่าจะเป็นของสัมประสิทธิ์ DCT เฉพาะ Coefficients-AC และการ fitting curve ด้วย Laplacian Distribution และ Cauchy Distribution ได้ดังนี้

○ : Akiyo sequence



รูปที่ 2 แสดง PDF ของ Coefficients -AC ของ Akiyo sequence ที่ตำแหน่ง (1,4)

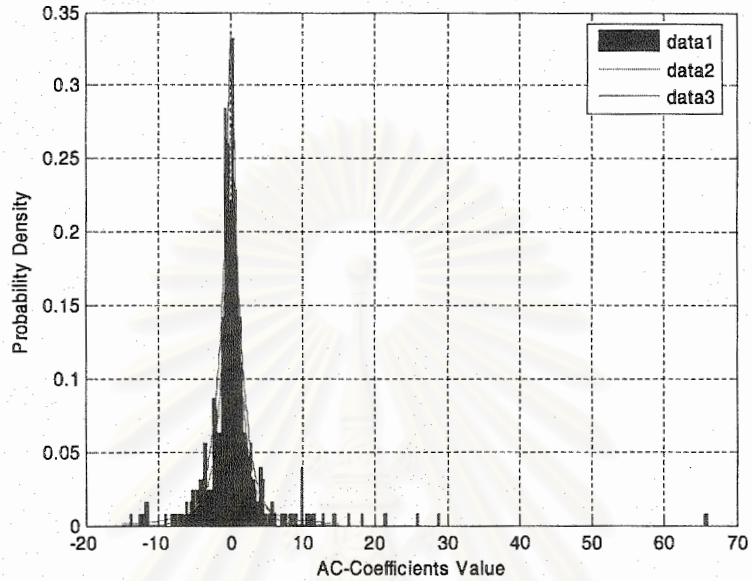
————— คือ : Cauchy Distribution $\lambda = 1.2 =$

————— คือ $1.9 = b$: Laplacian Distribution

	0	1	2	3	4	5	6	7
0	DC	8	4	2.5	1.8	1.5	1	0.6
1	6	3.2	2	1.5	1.2	1.2	0.65	0.45
2	2.8	1.9	1.4	1.4	1	0.9	0.6	0.5
3	2.3	1.4	1.1	1.2	0.9	0.75	0.6	0.5
4	1.6	1	0.9	0.9	0.7	0.6	0.65	0.45
5	1	0.9	0.9	0.9	0.7	0.7	0.75	0.35
6	0.8	0.8	0.6	0.6	0.5	0.45	0.4	0.35
7	0.55	0.45	0.45	0.45	0.45	0.3	0.25	0.25

รูปที่ 3 แสดง (Cauchy Parameter λ) ของ coefficients-AC ที่ตำแหน่งต่างๆ

○ : Carphone sequence



รูปที่ 4 แสดง PDF ของ Coefficients-AC ของ Carphone sequence ที่ตำแหน่ง (4,6)

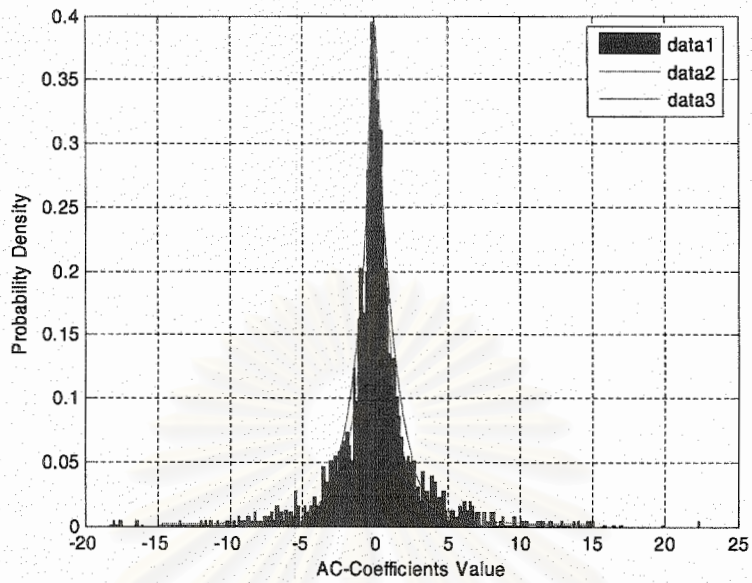
— คือ : Cauchy Distribution $\lambda_1 =$

— คือ $1.5 = b$: Laplacian Distribution

	0	1	2	3	4	5	6	7
0	DC	14	3.5	2.8	2.2	1.9	1.35	1.1
1	15.5	5.2	2.6	2.2	1.7	1.5	1.2	0.85
2	4.5	4.2	2.4	1.5	1.7	1.5	1.1	0.8
3	3.6	2.6	1.9	1.3	1.2	1.2	0.9	0.7
4	2.3	2.2	1.5	1.4	1.1	1.2	1	0.8
5	1.8	1.7	1.5	1.5	1.1	1.1	0.7	0.8
6	1.3	1.3	1.3	1	1	0.9	0.65	0.45
7	1.2	1.2	1.1	0.75	0.9	0.65	0.75	0.6

รูปที่ 5 แสดง (Cauchy Parameter λ) ของ coefficients-AC ที่ตำแหน่งต่างๆ

○ : Tempete Sequence



รูปที่ 6 แสดง PDF ของ Coefficients -AC ของ Tempete sequence ที่ตำแหน่ง (7,5)

———— คือ : Cauchy Distribution $\lambda 0.8 =$
 - - - - - คือ $1.25 = b$: Laplacian Distribution

	0	1	2	3	4	5	6	7
0	DC	17	10	10	6	6	4.5	3.5
1	17	11	8	5.5	5	4	3.5	3
2	13	8.5	6.5	5	4.5	3.5	3	2.5
3	9	6.5	4.5	4.5	3.2	3.2	2.5	2.2
4	6	4.5	4	3	2.5	2.5	1.8	1.5
5	4.3	3.5	2.5	2	1.8	1.7	1.2	1.1
6	2.6	2.2	1.8	1.5	1.2	1	0.9	0.7
7	1.7	1.5	1.3	1	0.9	0.8	0.65	0.55

รูปที่ 7 แสดง (Cauchy Parameter λ) ของ coefficients-AC ที่ตำแหน่งต่างๆ

2. พิจารณา Entropy function ซึ่งเกิดจากการ Quantized ค่าสัมประสิทธิ์ DCT ด้วย Uniform Quantier ดังสมการ

$$H(Q) = - \sum_{i=-\infty}^{\infty} P(iQ) \log_2 [P(iQ)]$$

โดยที่ความน่าจะเป็นของสัมประสิทธิ์ DCT ซึ่งจะถูก Quantized ด้วยค่า Q คือ

$$P(iQ) = \int_{(i-\frac{1}{2})Q}^{(i+\frac{1}{2})Q} f_X(x) dx$$

จากการทดลองในข้อ 1 จะเห็นได้ว่า PDF ของ coefficients-AC มีลักษณะใกล้เคียงกับ Cauchy Distribution มากกว่าแบบ Laplacian Distribution เนื่องจากเมื่อพิจารณาส่วนหางของ PDF จะพบว่า Cauchy Distribution สามารถครอบคลุมได้มากกว่าแบบ Laplacian Distribution

ดังนั้นค่า entropy ของ Cauchy Distribution จึงสามารถคำนวณหาได้จากสมการ

$$H(Q) = - \frac{2}{\pi} \tan^{-1} \left(\frac{Q}{2\mu} \right) \log_2 \left[\frac{2}{\pi} \tan^{-1} \left(\frac{Q}{2\mu} \right) \right] - \frac{2}{\pi} \sum_{i=1}^{\infty} \tan^{-1} \left(\frac{\mu Q}{\mu^2 + (i^2 - 1/4)Q^2} \right) \times \log_2 \left[\tan^{-1} \left(\frac{\mu Q}{\mu^2 + (i^2 - 1/4)Q^2} \right) \right]$$

(1) equation -----

3. เข้ารหัสสัญญาณวิดีโอด้วย Codec 264.H จำนวน 120 เฟรม เป็น (GOP) Group of picture [...IPPP] ขนาด 12 เฟรม สำหรับแต่ละ News, Foreman sequence, Carphone sequence, Akiyo) Sequence และ (Tempete sequence โดยปรับค่า Quantization parameter ตั้งแต่ [51-0] เพื่อดูค่า Actual rate (frame/bits) ของแต่ละ sequence แล้วทำการพล็อตกราฟเพื่อดูอัตราบิต โดยแบ่งเป็น Intra frame และ Nonintra frame

4. ศึกษาการทำงานของ การเข้ารหัสสัญญาณวิดีโอ ตามมาตรฐาน AVC/264.H ในส่วนของการ Transform และ Quantization

QP	0	1	2	3	4	5	6	7	8	9	10	11	12
QStep	0.625	0.6875	0.8125	0.875	1	1.125	1.25	1.375	1.625	1.75	2	2.25	2.5
QP	...	18	...	24	...	30	...	36	...	42	...	48	...	51
QStep		5		10		20		40		80		160		224

ตารางที่ 1. ค่า Quantization step sizes ใน Codec 264.H

จากตารางแสดงให้เห็นว่า ค่า Quantization parameter ที่มีค่าอยู่ในช่วง [51-0] เป็นเพียงตัวเลขดัชนี (Index) ที่ใช้บอกค่า Quantization step sizes เท่านั้น ซึ่ง Quantization step sizes จะมีค่าเพิ่มขึ้น 2 เท่า ที่

Quantization parameter มีการเพิ่มขึ้นทุกๆ 6 ระดับ นั่นคือ Quantization step sizes จะมีค่าเพิ่มขึ้นประมาณ % 12.5 สำหรับการเพิ่มขึ้นของ 1r Quantization parameter ระดับ

ดังนั้นการคำนวณหาค่า (pixel/bits)Entropy จากสมการ Entropy Cauchy function โดยใช้ค่า Quantization parameter จึงไม่ถูกต้อง เพราะการเข้ารหัสจริงใน Codec 264.H นั้นค่า Transform coefficients จะถูก Quantized ด้วยค่า Quantization step sizes ด้วยสมการ

$$Z_{ij} = \text{round} \left[\frac{Y_{ij}}{Q_{step}} \right]$$

โดยที่ Z_{ij} คือ Coefficients Quantization
 Y_{ij} คือ Transform coefficients
 Q_{step} คือ zesQuantization step si

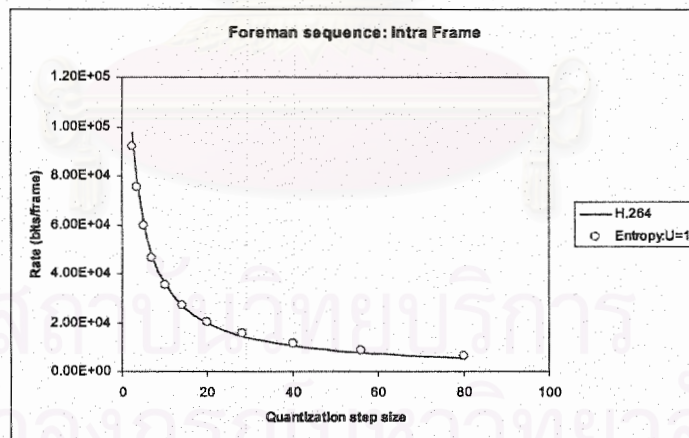
ดังนั้นค่า Q ในสมการที่ (1) จึงหมายถึงค่า Quantization step size แทนที่จะเป็นค่า Quantization parameter .5 คำนวณหาค่าจำนวนหาค่า (pixel/bits)Entropy จากสมการ) Entropy Cauchy function สมการที่ ((1) ใหม่ที่ค่า λ) ในรูปใช้ค่า U แทน (โดยใช้ค่า ion step sizes Quantizat ต่างๆ ดังในตารางที่ .1 แล้วทำการพล็อตกราฟอัตราบิต (frame/bits) เปรียบเทียบระหว่าง Actual rate กับ ค่าที่คำนวณได้จากสมการที่ .1

$$= (\text{frame/bits})\text{Rate} \quad H(Q) \times (144 \times 176 \times 1.5) \quad \text{สำหรับกรณีที่กำลังภาพเป็นชนิด QCIF}$$

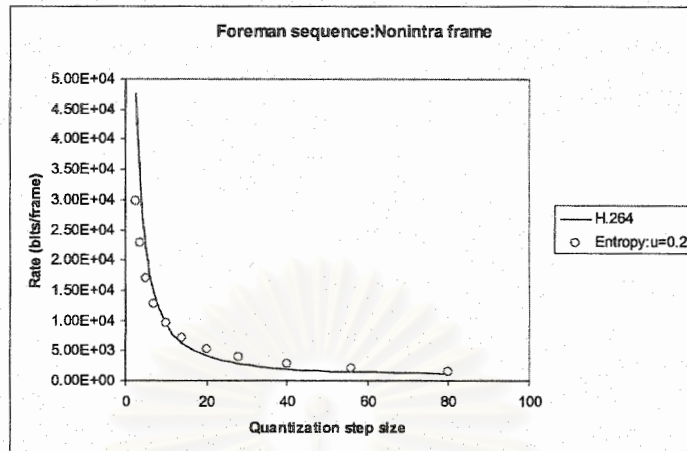
$$= (\text{ramef/bits})\text{Rate} \quad H(Q) \times (288 \times 352 \times 1.5) \quad \text{สำหรับกรณีที่กำลังภาพเป็นชนิด CIF}$$

ซึ่งสามารถแสดงผลได้ดังนี้

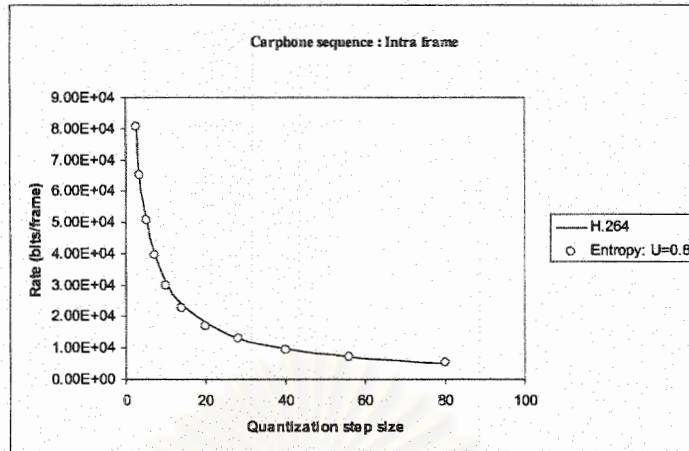
- ลำดับภาพ foreman



(a)

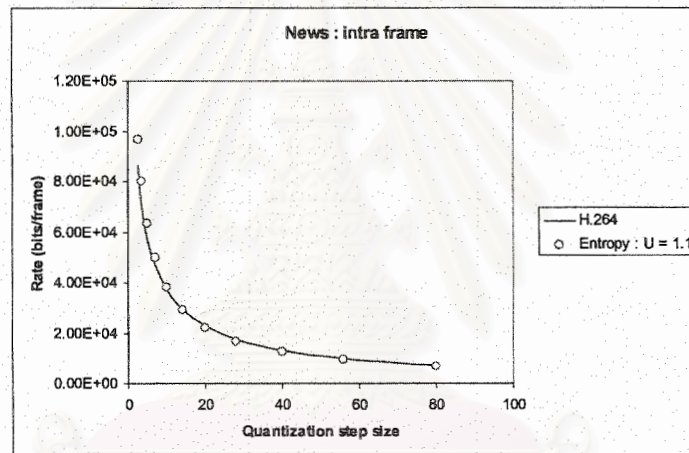


- ลำดับภาพ Carphone



รูปที่ 10 แสดงความสัมพันธ์ระหว่างอัตราบิตกับค่าควอนไทซ์ต่างๆ ของลำดับภาพ Carphone

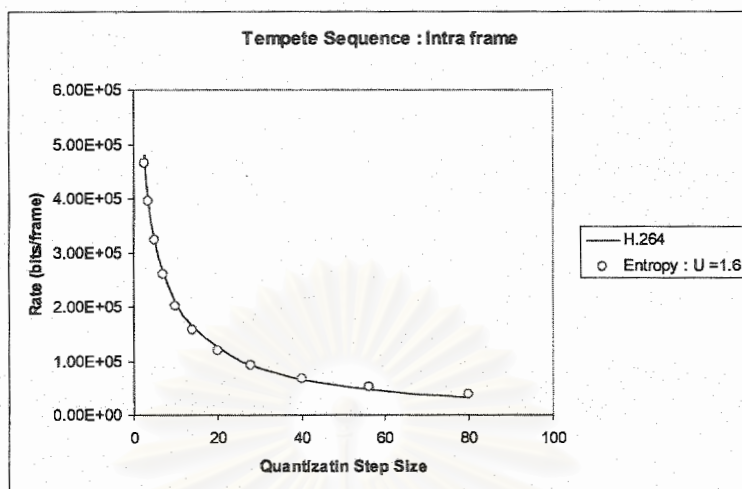
- ลำดับภาพ News



รูปที่ 11 แสดงความสัมพันธ์ระหว่างอัตราบิตกับค่าควอนไทซ์ต่างๆ ของลำดับภาพ ewsN

- ลำดับภาพ Tempete

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 12 แสดงความสัมพันธ์ระหว่างอัตราบิตกับค่าควอนไทซ์ต่างๆ ของลำดับภาพ Tempete

จากกราฟ แสดงให้เห็นได้ว่าอัตราบิต (frame/bits) ที่คำนวณได้จากสมการ Cauchy Entropy function กับอัตราบิตที่ได้จากการเข้ารหัสจริง (Actual rate) ด้วย 264.H นั้นมีความสอดคล้องกัน ดังนั้นสมการ Cauchy Entropy function จึงสามารถนำมาใช้ในการคำนวณ Rate model สำหรับการเข้ารหัสได้

ปรับปรุงการเข้ารหัสสัญญาณวิดีโอในส่วนของกระบวนการควบคุมอัตราในมาตรฐาน AVC/264.H โดยใช้

Distortion Model-Cauchy Rate

1. วิเคราะห์สมการทางคณิตศาสตร์เพื่อหาความสัมพันธ์ระหว่างแบบจำลองอัตราบิต และอัตราความเพี้ยนที่เหมาะสม โดยใช้ Lagrange Multiplier Technique

จากบทความเรื่อง “Density-AVC video coder via Cauchy/264.Frame bit allocation for the H-Based rate and Distortion models” สามารถคำนวณหา Model-Rate และ Distortion Model ได้จากสมการ (2) และ (3) ตามลำดับ

$$R = H(Q) = aQ^{-\alpha} \quad (2)$$

$$D(Q) = bQ^\beta \quad (3)$$

โดยที่ a, α, b, β คือค่าพารามิเตอร์ของแบบจำลอง

กำหนดให้ C = จำนวนพิกเซลใน basic unit

(pixel/itsb) Average rate = H สำหรับการเข้ารหัสเวกเตอร์การเคลื่อนที่ และ coder's

.header bits

ดังนั้นจาก ; (2)

$$R_i = Ca_i Q_i^{-\alpha_i} + H_i \quad (4)$$

โดยที่ โดยที่ $i = 1, 2, \dots, N$ และ $N =$ จำนวน basic unit ภายในเฟรม จาก (3) หา Distortion ของ ic unit bas ที่ i จาก

$$D = \frac{1}{N} \sum_{i=1}^N b_i Q_i^{\beta_i} \quad (5)$$

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

$$J(Q_1^*, Q_2^*, \dots, Q_N^*, \lambda^*) = \frac{1}{N} \sum_{i=1}^N b_i Q_i^{\beta_i} + \lambda \sum_{i=1}^N (Ca_i Q_i^{-\alpha_i} + H_i) - \lambda R_{MAX}$$

$$\therefore J(Q_1^*, Q_2^*, \dots, Q_N^*, \lambda^*) = \frac{1}{N} \sum_{i=1}^N b_i Q_i^{\beta_i} + \lambda \sum_{i=1}^N (Ca_i Q_i^{-\alpha_i}) + \lambda \sum_{i=1}^N H_i - \lambda R_{MAX} \quad (6)$$

โดยที่ Constraints bit rate = R

หาค่า Q_N^* และ λ^* จากการทำ Partial Derivative สมการ (6)

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

$$\frac{\partial J}{\partial Q_1^*} = \frac{\beta_1}{N} b_1 (Q_1^*)^{(\beta_1-1)} - (\lambda Ca_1 \alpha_1) (Q_1^*)^{-(\alpha_1+1)}$$

$$\frac{\partial J}{\partial Q_2^*} = \frac{\beta_2}{N} b_2 (Q_2^*)^{(\beta_2-1)} - (\lambda Ca_2 \alpha_2) (Q_2^*)^{-(\alpha_2+1)}$$

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

$$\therefore \frac{\partial J}{\partial Q_i^*} = \left(\frac{\beta_i b_i}{N} \right) (Q_i^*)^{(\beta_i-1)} - (\lambda Ca_i \alpha_i) (Q_i^*)^{-(\alpha_i+1)} \quad (7)$$

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

$$\therefore \frac{\partial J}{\partial \lambda^*} = C \sum_{i=1}^N a_i (Q_i^*)^{-\alpha_i} + (NH_i - R_{MAX}) \quad (8)$$

จาก (7) กำหนดให้ $\frac{\partial J}{\partial Q_i^*} = 0$

ดังนั้น
$$\left(\frac{\beta_i b_i}{N}\right) (Q_i^*)^{(\beta_i-1)} = (\lambda C a_i \alpha_i) (Q_i^*)^{-(\alpha_i+1)}$$

$$\therefore (Q_i^*)^{(\beta_i+\alpha_i)} = \left(\frac{\lambda N C a_i \alpha_i}{\beta_i b_i}\right) \tag{9}$$

จาก ; (8) กำหนดให้ $\frac{\partial J}{\partial \lambda^*} = 0$

$$\therefore \sum_{i=1}^N a_i (Q_i^*)^{-\alpha_i} = \frac{(R_{MAX} - NH_1)}{C} \tag{10}$$

พิจารณา Take ln , (9) ทั้งสองข้างของสมการ

$$\ln(Q_i^*)^{(\beta_i+\alpha_i)} = \ln\left(\frac{\lambda N C a_i \alpha_i}{\beta_i b_i}\right)$$

$$\ln(Q_i^*) = \frac{1}{(\beta_i + \alpha_i)} \ln\left(\frac{\lambda N C a_i \alpha_i}{\beta_i b_i}\right)$$

$$\therefore Q_i^* = \left(\frac{\lambda N C a_i \alpha_i}{\beta_i b_i}\right)^{\frac{1}{(\beta_i + \alpha_i)}} \tag{11}$$

นำ (11) แทนใน ; (10)

$$\sum_{i=1}^N a_i \left(\frac{\lambda N C a_i \alpha_i}{\beta_i b_i}\right)^{\frac{-\alpha_i}{(\beta_i + \alpha_i)}} = \frac{(R_{MAX} - NH_1)}{C}$$

$$\therefore \lambda = \left(\frac{R_{MAX} - NH_1}{N C^2}\right) \left(\frac{1}{\sum_{i=1}^N a_i \left(\frac{a_i \alpha_i}{\beta_i b_i}\right)^{\frac{-\alpha_i}{(\beta_i + \alpha_i)}}}\right) \tag{12}$$

นำ (12) แทนใน ; (11)

$$\therefore Q_i^* = \left\{ \left(\frac{R_{MAX} - NH_i}{C} \right) \cdot \left(\frac{1}{\sum_{i=1}^N a_i \left(\frac{a_i \alpha_i}{\beta_i b_i} \right)^{\frac{-\alpha_i}{(\beta_i + \alpha_i)}}} \right) \cdot \left(\frac{a_i \alpha_i}{\beta_i b_i} \right) \right\}^{\frac{1}{(\beta_i + \alpha_i)}} \quad (13)$$

2. ทำการ Implement ในส่วนของ Rate control ลงในมาตรฐานการเข้ารหัสสัญญาณวิดีโอ H.264 โดย ใช้ Distortion model-Cauchy Rate และการใช้ Quadratic rate model ซึ่งเป็นแบบจำลองอัตราบิตที่ได้จากการวิเคราะห์ PDF ของ Coefficients-AC ของลำดับภาพเป็นแบบ Laplacian distribution โดยแบ่งขั้นตอนการทำงานออกเป็น GOP-Layer-Frame, Layer และ Basic Unit-Layer ดังนี้

- การควบคุมอัตราในระดับ : Layer GOP คำนวณหาจำนวนบิตรวมของเฟรมที่เหลืออยู่ใน GOP $B_i(j)$ (เมื่อเฟรมที่ j ของ GOP ที่ i ถูกเข้ารหัส

$$T_r(n_{i,j}) = \begin{cases} \frac{u(n_{i,1})}{F_r} \times N_{gop} - B_c(n_{i-1, N_{gop}}) & 1 = j ; \\ T_r(n_{i,j-1}) - b(n_{i,j-1}) & j = 2, 3, \dots, N \end{cases} \quad (14)$$

- ควบคุมอัตราในระดับ : Frame Layer คำนวณหาจำนวนบิตเป้าหมายของแต่ละเฟรม คำนวณหาค่าควอนไทซ์เริ่มต้นสำหรับ I เฟรม และ P เฟรมแรก Q_i^0, Q_p^0 (โดยพิจารณาจาก Channel bandwidth และ ขนาดของ) GOP ตามมาตรฐาน (264.H และ คำนวณหาอัตราบิตเป้าหมายของแต่ละ P เฟรมดังนี้

(1 คำนวณหาระดับบัพเฟอร์เป้าหมาย $Tbl(n_{i,j+1})$ ของแต่ละ P เฟรมจากสมการ

$$Tbl(n_{i,j+1}) = Tbl(n_{i,j}) - \frac{Tbl(n_{i,2})}{N_p - 1} \quad (15)$$

(2 คำนวณหาจำนวนบิตเป้าหมายของแต่ละ P เฟรม โดยพิจารณาจาก 2 กรณีคือ

- พิจารณาจากระดับบัพเฟอร์เป้าหมาย ซึ่งจำนวนบิตเป้าหมายสามารถคำนวณได้จากสมการ

$$\tilde{f}(n_{i,j}) = \frac{u(n_{i,j})}{F_r} + \gamma (Tbl(n_{i,j}) - B_c(n_{i,j})) \quad (16)$$

- พิจารณาจากจำนวนบิตที่เหลืออยู่ของแต่ละเฟรม

$$\hat{f}(n_{i,j}) = \frac{T_r(n_{i,j})}{N_p - j} \quad (17)$$

ซึ่งในการหาจำนวนบิตเป้าหมายของแต่ละ P เฟรมนั้นจะพิจารณาจาก ทั้ง 2 กรณี คือ

$$f(n_{i,j}) = \beta \hat{f}(n_{i,j}) + (1 - \beta) \tilde{f}(n_{i,j}) \quad (18)$$

โดยที่ β เป็นค่าถ่วงน้ำหนัก โดยกำหนดให้มีค่าเท่ากับ 0.5

- ควบคุมอัตราในระดับ : Basicunit Layerแบ่งออกเป็น 2 ขั้นตอนคือ

1) : encoding stage-Preคำนวณหาค่าระดับขั้นของการควอนไทซ์สำหรับแต่ละหน่วยย่อยพื้นฐาน

กรณี 1 สำหรับหน่วยย่อยพื้นฐานแรกของแต่ละ Pเฟรม ค่าควอนไทซ์เซชันพารามิเตอร์จะพิจารณาจากค่าควอนไทซ์เซชันพารามิเตอร์ของทุกๆหน่วยย่อยพื้นฐานใน Pเฟรมก่อนหน้า

$$QP_{1,i}(j) = \text{Avg}QP_1(j-1) \quad (19)$$

กรณี 2 เมื่อจำนวนบิตที่เหลืออยู่มีค่าน้อยกว่าศูนย์ ค่าควอนไทซ์เซชันพารามิเตอร์จำกัดจำนวนจาก

$$QP_{1,i}(j) = QP_{1,i}(j) + DQUANT \quad (20)$$

โดยที่ $DQUANT = \begin{cases} 1, QP_{1,i}(j) > 25 \\ 2, QP_{1,i}(j) \neq 25 \end{cases}$

กรณี 3 คำนวณหาค่าระดับขั้นของการควอนไทซ์จาก Cauchy Rate-Distortion Model

$$Q_l(n_{i,j}) = \left\{ \left(\frac{f(n_{i,j}) - Nm_{hdr,l}}{C} \right) \cdot \left(\frac{1}{\sum_{l=1}^N a_{j,l} \left(\frac{a_{j,l} \alpha_{j,l}}{\beta_{j,l} b_{j,l}} \right)^{\frac{-\alpha_{j,l}}{(\beta_{j,l} + \alpha_{j,l})}}} \right) \cdot \left(\frac{a_{j,l} \alpha_{j,l}}{\beta_{j,l} b_{j,l}} \right) \right\}^{\frac{1}{(\beta_{j,l} + \alpha_{j,l})}} \quad (21)$$

Post Encoding Stage (2)

Update - ค่าพารามิเตอร์โมเดล a และ α โดยพิจารณาจาก

$$\ln(R_l(n_{i,j})) = \ln(a_{j,l}) - \alpha_{j,l} \ln(Q_l(n_{i,j})) \quad (22)$$

จากการใช้ โดยใช้ Linear regression analysisจะสามารถคำนวณค่าพารามิเตอร์โมเดลได้จากสมการ

$$\ln(a_{j,l}) = \frac{\left(\sum_k^{l-1} \ln(R_k) \right) \left(\sum_k^{l-1} (\ln(Q_k))^2 \right) - \left(\sum_k^{l-1} \ln(Q_k) \ln(R_k) \right)}{(l-k) \left(\sum_k^{l-1} (\ln(Q_k))^2 \right) \left(\sum_k^{l-1} \ln(Q_k) \right)^2} \quad (23)$$

$$\alpha_{j,l} = \frac{\left(\sum_k^{l-1} \ln(Q_k) \right) \left(\sum_k^{l-1} (\ln(R_k)) \right) - (l-k) \left(\sum_k^{l-1} \ln(Q_k) \ln(R_k) \right)}{(l-k) \left(\sum_k^{l-1} (\ln(Q_k))^2 \right) \left(\sum_k^{l-1} \ln(Q_k) \right)^2} \quad (24)$$

Update-ค่าพารามิเตอร์โมเดล b และ β โดยพิจารณาจาก

$$\ln(MSE_l(n_{i,j})) = \ln(b_{j,l}) + \beta_{j,l} \ln(Q_l(n_{i,j})) \quad (25)$$

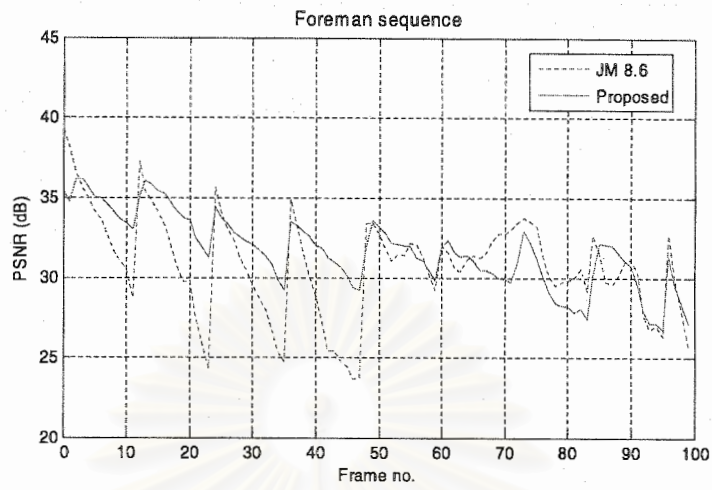
จากการใช้ โดยใช้ Linear regression analysis จะสามารถคำนวณหาค่าพารามิเตอร์โมเดลได้จากสมการ

$$\ln(b_{j,l}) = \frac{\left(\sum_k^{l-1} \ln(MSE_k) \right) \left(\sum_k^{l-1} (\ln(Q_k))^2 \right) - \left(\sum_k^{l-1} \ln(Q_k) \ln(MSE_k) \right)}{(l-k) \left(\sum_k^{l-1} (\ln(Q_k))^2 \right) \left(\sum_k^{l-1} \ln(Q_k) \right)^2} \quad (26)$$

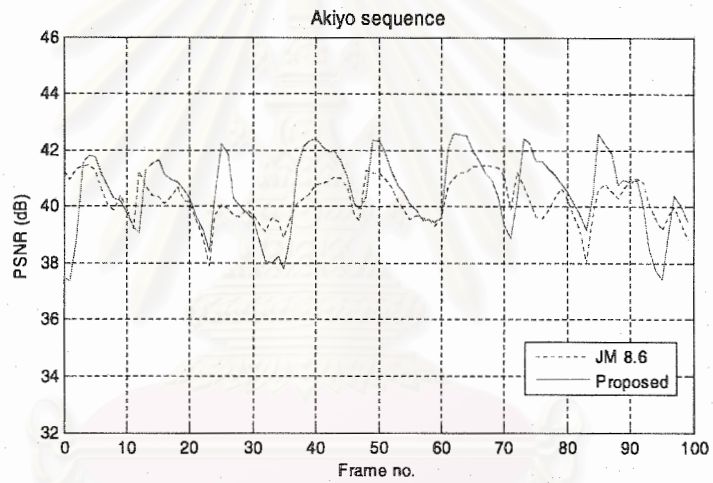
$$\beta_{j,l} = \frac{(l-k) \left(\sum_k^{l-1} \ln(Q_k) \ln(MSE_k) \right) - \left(\sum_k^{l-1} \ln(Q_k) \right) \left(\sum_k^{l-1} \ln(MSE_k) \right)}{(l-k) \left(\sum_k^{l-1} (\ln(Q_k))^2 \right) \left(\sum_k^{l-1} \ln(Q_k) \right)^2} \quad (27)$$

3. สรุปผลการทดลอง

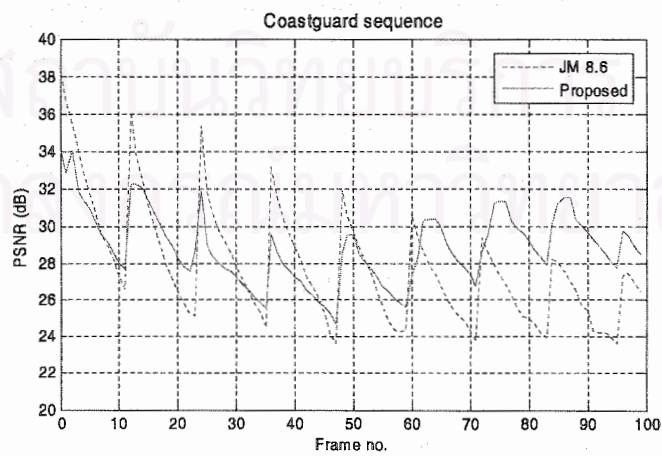
ทำการทดลองใช้ Cauchy Rate Distortion Model ที่วิเคราะห์ได้ ตามรายละเอียดที่ผ่านมาตรฐาน (8.6JM) AVC /264.H โดยทำการเข้ารหัสแบบ PL..GOP IP โดย GOP 1 ประกอบด้วย -I 1 เฟรม และ -P 11 เฟรม เข้ารหัสที่อัตราบิตต่างๆ คือ kbps 128 , kbps 64 , kbps 32 , kbps 16 และ kbps 256 ใช้ว่า คับภาพทดสอบจำนวน 5 ลำดับภาพ (Foreman, News, Akiyo, Carphone และ Coastguard (ซึ่งจากผลการทดลองพบว่า Cauchy Rate Distortion Model ที่นำเสนอ นั้นสามารถปรับปรุงค่า PSNR ได้สูงขึ้นเมื่อเทียบกับการเข้ารหัสตามมาตรฐาน 264.H โดยสามารถปรับปรุงค่า PSNR ได้สูงขึ้น , dB 0.65 0.68 dB 0.73 , dB 0.74 , dB และ dB 1.18 ที่อัตราบิต kbps 128 , kbps 64 , kbps 32 , kbps 16 และ kbps 256 ตามลำดับ ซึ่งสามารถแสดงรายละเอียดของผลการทดลองได้ในตารางที่ 2 ส่วนกราฟแสดงความสัมพันธ์ระหว่างค่า PSNR ของแต่ละเฟรม นั้นสามารถแสดงได้ในรูปที่ 13 และ 14



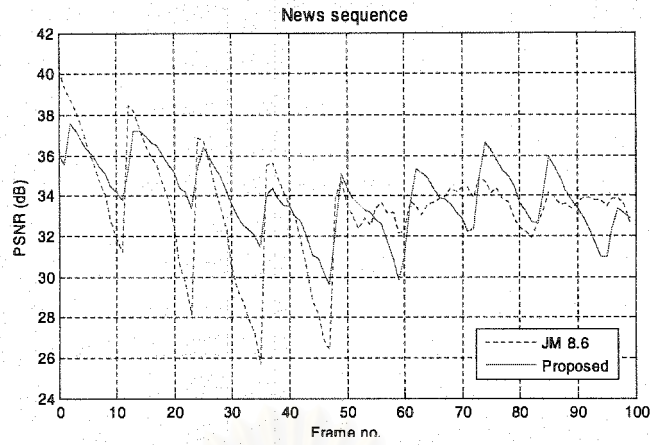
(a)



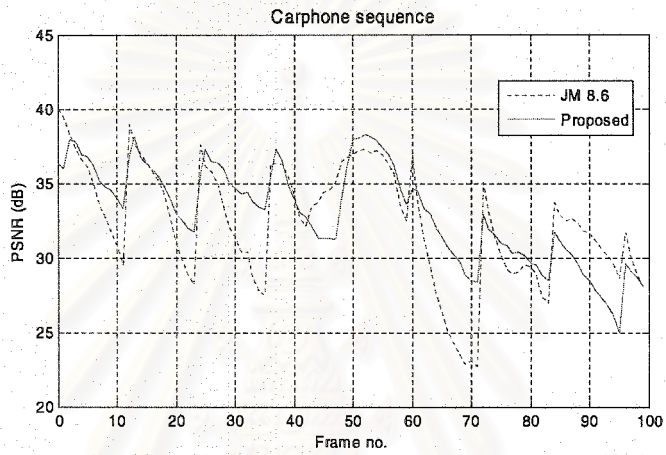
(b)



(c)



(d)

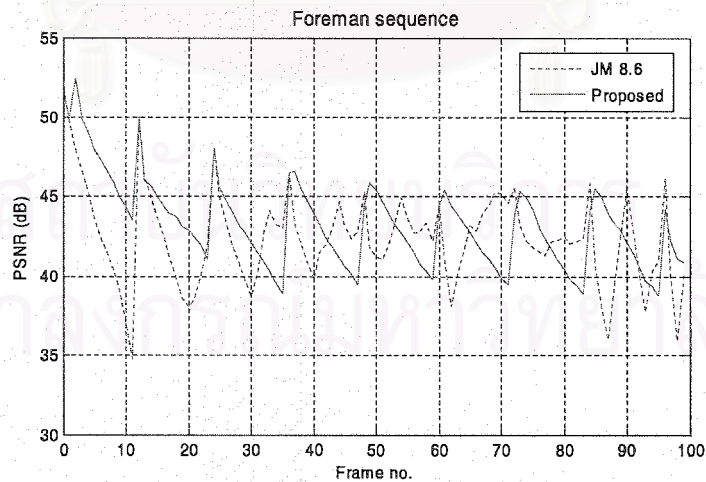


(e)

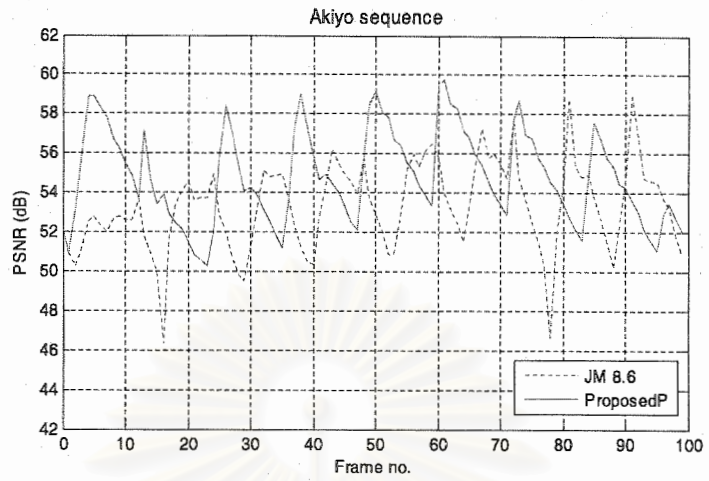
รูปที่ 13 แสดงค่า PSNR ของแต่ละเฟรม สำหรับแต่ละลำดับภาพทดสอบที่อัตราบิต kbps 32

(a) foreman sequence (b) Akiyo sequence (c) Coastguard sequence

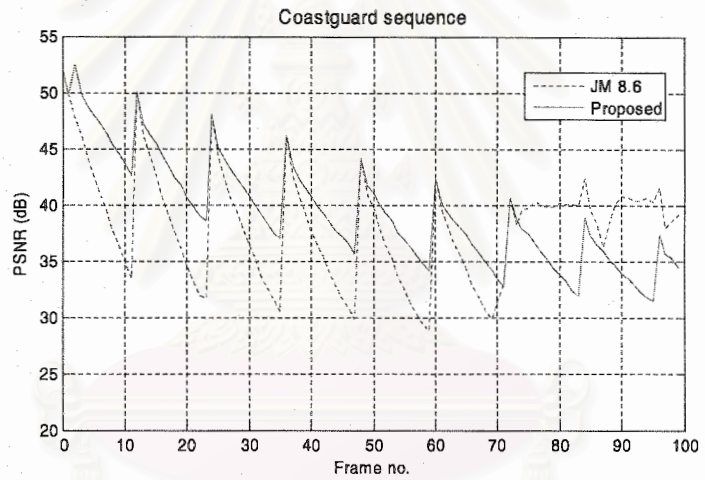
(d) News sequence and (e) Carphone sequence



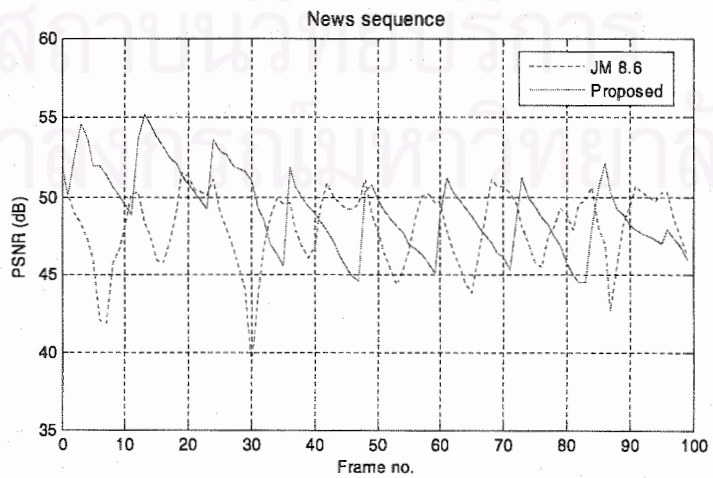
(a)



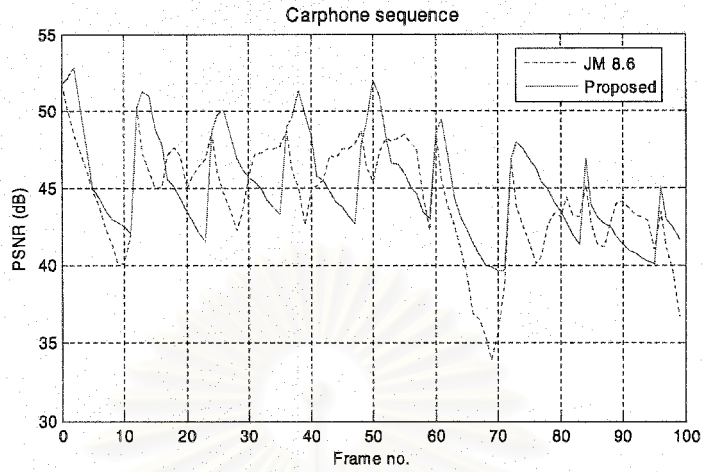
(b)



(c)



(d)

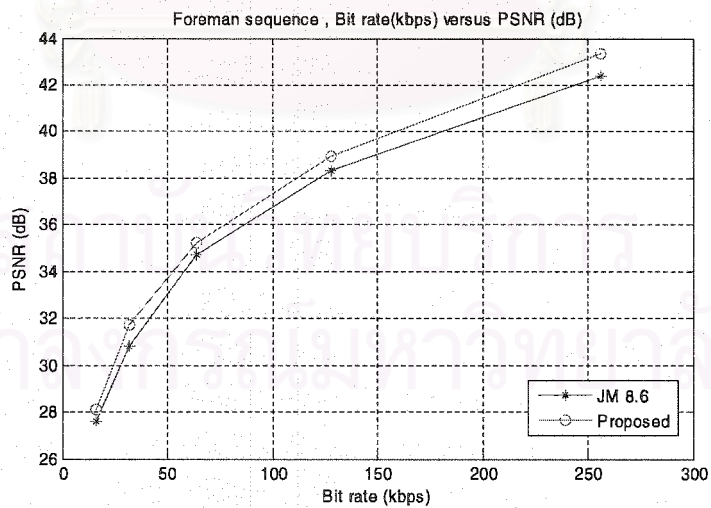


(e)

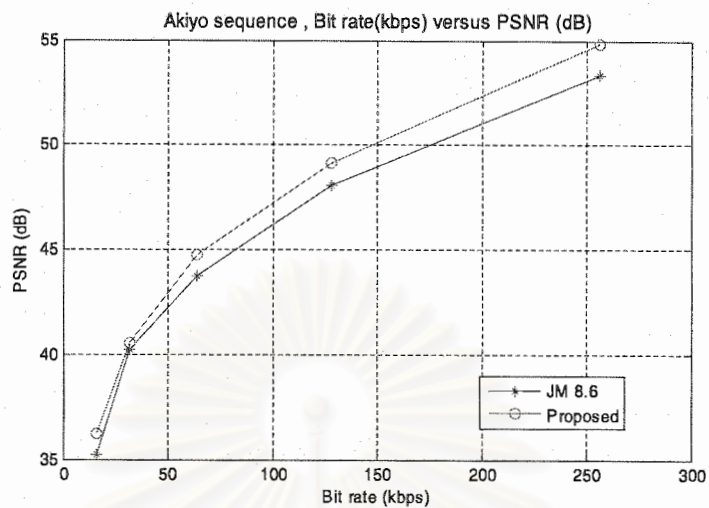
รูปที่ 14 แสดงค่า PSNR ของแต่ละเฟรม สำหรับแต่ละลำดับภาพทดสอบที่อัตราบิต kbps 256

- (a) foreman sequence (b) Akiyo sequence (c) Coastguard sequence
(d) News sequence and (e) Carphone sequence

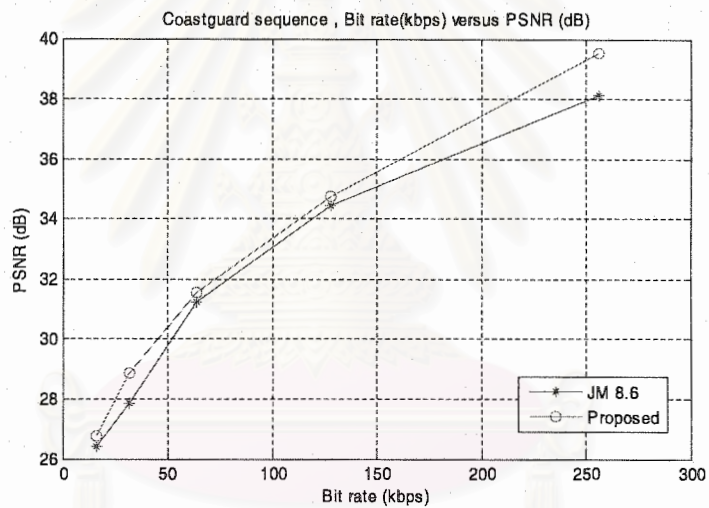
รูปที่ 15 แสดงความสัมพันธ์ระหว่างค่า PSNR ของทุกๆลำดับภาพทดสอบ กับอัตราบิตต่างๆ ซึ่งจากรูปแสดงให้เห็นว่างานวิจัยสามารถปรับปรุงค่า PSNR ได้สูงขึ้นที่ทุกๆ อัตราบิตเมื่อเทียบกับมาตรฐาน AVC/264.H.264 เดิม



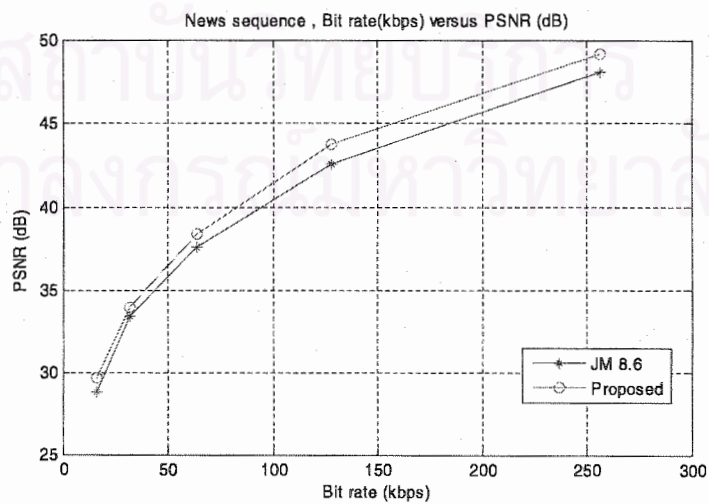
(a)



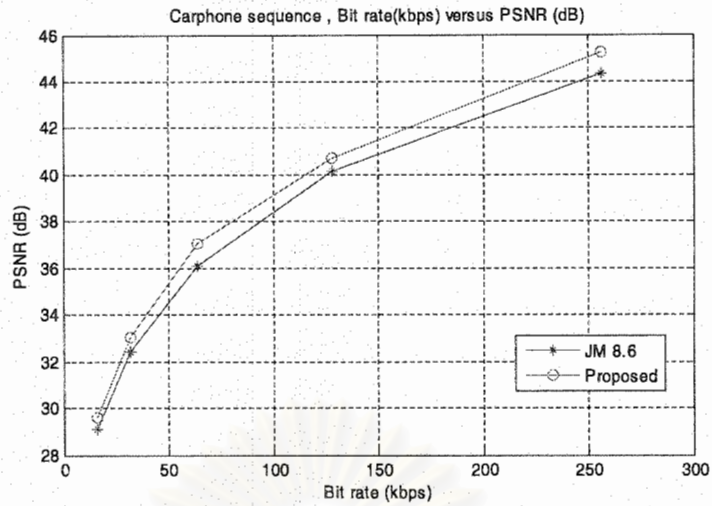
(b)



(c)



(d)



(c)

รูปที่ 15 แสดงค่า PSNR แต่ละลำดับภาพทดสอบที่อัตราบิตต่างๆ

(a) foreman sequence (b) Akiyo sequence (c) Coastguard sequence

(d) News sequence and (e) Carphone sequence

ตารางที่ 2 สรุปผลการทดลองการเข้ารหัสสัญญาณวิดีโอที่อัตราบิตต่างๆ

Bit rate	quenceSe	(dB)Average PSNR		. PSNR Std		(kbps)Bit Rate	
		8.6JM	Proposed	8.6JM	Proposed	8.6JM	Proposed
kbps 16	Foreman	27.58	0.53+) 28.11 (dB)	1.81	1.70	16.50	16.25
	Coastguard	26.43	0.35+) 26.78 (dB)	1.30	1.21	16.12	16.56
	Carphone	29.10	0.53+) 29.63 (dB)	70.2	2.70	16.26	16.20
	Akiyo	35.20	(dB 1.0+) 36.20	1.54	1.42	16.22	16.35
	News	28.83	0.86+) 29.69 (dB)	1.42	1.13	16.17	16.15
kbps 32	Foreman	30.81	0.90+) 31.71 (dB)	3.16	2.27	32.34	32.39
	Coastguard	27.88	0.99+) 28.87 (dB)	3.17	1.94	32.34	32.52

	Carphone	32.42	0.64+) 33.06 (dB	3.93	3.23	32.88	32.74
	Akiyo	40.24	0.28+) 40.52 (dB	0.79	1.34	33.16	33.02
	News	33.42	0.58+) 34.00 (dB	2.66	1.78	32.94	33.04
kbps 64	Foreman	34.73	0.52+) 35.25 (dB	1.72	1.72	64.80	65.01
	Coastguard	31.23	0.34+) 31.57 (dB	2.20	07.2	65.97	65.48
	Carphone	36.11	0.96+) 37.07 (dB	3.45	3.11	64.73	64.56
	Akiyo	43.72	1.06+) 44.78 (dB	1.90	2.01	65.42	65.49
	News	37.67	0.79+) 38.46 (dB	1.76	1.71	64.76	65.02
kbps 128	Foreman	38.31	0.65+) 38.96 (dB	2.04	2.35	127.67	127.78
	Coastguard	46.34	0.28+) 34.74 (dB	2.70	2.85	127.77	128.01
	Carphone	40.14	0.61+) 40.75 (dB	3.56	3.49	129.32	128.95
	Akiyo	48.10	1.06+) 49.16 (dB	1.90	1.81	129.21	129.56
	News	42.63	1.11+) 43.74 (dB	1.97	1.95	129.02	128.45
kbps 256	Foreman	42.40	dB 0.97+) 43.37 (01.3	2.91	255.92	256.31
	Coastguard	38.12	(dB 1.4+) 39.52	4.90	4.90	257.59	257.58
	Carphone	44.36	0.86+) 45.22 (dB	3.54	3.31	256.01	256.12
	Akiyo	53.31	54.84+) 1.53	2.22	2.35	256.98	256.14

			(dB				
	News	48.06	1.13+) 49.19 (dB	2.42	2.50	257.54	257.17

4. งานที่จะทำต่อไป

4.1 ศึกษาคุณลักษณะของการส่งสัญญาณวิถีทัศน์ผ่านช่องสัญญาณไร้สาย กรรมวิธีแก้ไขความผิดพลาดแบบต่างๆ (ARQ, FEC)

4.2 หาแบบจำลองความเพี้ยนเนื่องมาจากการส่งสัญญาณวิถีทัศน์ผ่านช่องสัญญาณไร้สาย และความเพี้ยนเนื่องมาจากการเข้ารหัสสัญญาณ

4.3 นำเอาแบบจำลอง R-D มาใช้ร่วมกับการควบคุมอัตรา โดยพิจารณาร่วมกับสถานะของช่องสัญญาณไร้สาย และแบบจำลองความเพี้ยน เพื่อใช้ในการหาค่าควอนไทซ์เซชันที่เหมาะสม เพื่อให้การเข้ารหัส และการส่งผ่านช่องสัญญาณไร้สายมีประสิทธิภาพสูงสุด

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย