บทที่ 2

วิธีลดปริมาณข้อมูลของภาพเคลื่อนไหว

ในบทนี้จะกล่าวถึงลักษณะทั่วไปของภาพเคลื่อนไหว วิธีการต่าง ๆ ที่ใช้ในการลดข้อมูลของภาพเคลื่อนไหว และแนวทางในการออกแบบระบบส่งภาพเคลื่อนไหวในวิทยานิพนธ์นี้

2.1 ลักษณะทั่วไปของภาพเคลื่อนไหว

สัญญาณภาพเคลื่อนไหวที่เราคุ้นเคยมากที่สุดคือสัญญาณโทรทัศน์ สัญญาณโทรทัศน์จัดว่าเป็นสัญญาณภาพเคลื่อนไหวที่มีคุณภาพดี เพราะได้คำนึงถึงความรู้สึกตอบสนองของมนุษย์ไว้เป็นอย่างดี กล่าวคือในแต่ละภาพจะประกอบด้วยเส้นแนวนอน 525 เส้น (ระบบ NTSC) และในหนึ่งวินาทีจะมีการเปลี่ยนแปลงของภาพ 30 ภาพ ลักษณะดังกล่าวนี้จะทำให้ตามนุษย์สามารถรับชมภาพได้โดยมีความรู้สึกเป็นธรรมชาติ คือการเคลื่อนไหวของภาพจะต่อเนื่องเป็นธรรมชาติ และความละเอียดของภาพ (525 เส้นต่อหนึ่งภาพ) ก็อยู่ในระดับที่ละเอียดเพียงพอ สัญญาณโทรทัศน์ที่มีคุณสมบัติดังกล่าวนี้จะมีแบนด์วิดท์ประมาณ 4 MHz โดยมีความถี่ต่ำสุดเป็น 50 Hz และความถี่สูงสุดประมาณ 4 MHz โดยที่ความถี่ต่ำสุดถูกกำหนดจากสัญญาณซิงค์ในแนวตั้ง และความถี่สูงสุดถูกกำหนดจากความละเอียดของภาพ

ในการส่งสัญญาณโทรทัศน์ออกไปตามช่องสัญญาณนั้นถ้าใช้วิธีมอดูเลตแบบ FM ซึ่งจะได้สัญญาณที่มีคุณภาพดีจะต้องใช้แบนด์วิดท์ของช่องสัญญาณมากกว่า 4 เท่าตัวขึ้นไป คือต้องใช้แบนด์วิดท์มากกว่า 16 MHz ซึ่งนับว่าค่อนข้างจะสิ้นเปลือง ดังนั้นในระบบกระจายสัญญาณโทรทัศน์ตามมาตรฐานทั่วไปจึงใช้การมอดูเลตแบบ Vestigial Sideband ซึ่งต้องการแบนด์วิดท์ประมาณ 7 MHz อย่างไรก็ตามเมื่อต้องการส่งทอด (relay) สัญญาณโทรทัศน์ไปยังที่ห่างไกลจำเป็นต้องใช้ระบบมอดูเลตที่มีคุณภาพดี เพราะฉะนั้นในสมัยเริ่มแรกที่ใช้ระบบแอนาลอกนั้นการส่งทอดสัญญาณโทรทัศน์จะใช้ระบบมอดูเลตแบบ FM ในระยะ

ต่อมาระบบส่งผ่านสัญญาณในระบบโทรศัพท์ได้เปลี่ยนมาเป็นการส่งแบบดิจิตอล ทำให้การส่งทอดสัญญาณโทรทัศน์เปลี่ยนมาเป็นการส่งแบบดิจิตอลเช่นเดียวกัน การส่งสัญญาณแอนาลอกโดยแปลงเป็นสัญญาณดิจิตอลแล้วค่อยส่งนั้นจะใช้หลักการของทฤษฎีสุ่มตัวอย่างซึ่งกล่าวไว้ว่าจะต้องสุ่มตัวอย่างด้วยความถี่สูงกว่า 2 เท่าของความถี่สูงสุดเพื่อให้สามารถแปลงสัญญาณดิจิตอลนั้นกลับมาเป็นสัญญาณแอนาลอกที่ถูกต้องเหมือนเดิมได้ จากนั้นถ้าทำการมอดูเลตแบบ PCM (pulse code modulation) ก็จะต้องใช้จำนวนบิตไม่น้อยกว่า 8 บิต (256 ระดับ) ต่อหนึ่งสัญญาณสุ่มตัวอย่าง นั่นคืออัตราความเร็วของสัญญาณดิจิตอลจะต้องสูงกว่า 16 เท่าของความถี่สูงสุดของสัญญาณโทรทัศน์ คือต้องส่งด้วยอัตราความเร็วสูงกว่า 64 Mbps อัตราความเร็วดังกล่าวนี้จัดว่าสูงมากและมีต้นทุนสูงจึงไม่เหมาะที่จะนำมาใช้ในการให้บริการโทรศัพท์ภาพหรือการประชุมทางไกล (teleconference) เพราะจะทำให้ค่าบริการสูงมาก

จากที่กล่าวมาข้างต้นจะเห็นได้ว่าการให้บริการโทรศัพท์ภาพ และการประชุมทางไกลโดยผ่านเข้าไปในระบบโทรศัพท์นั้นมีความจำเป็นอย่างยิ่งที่จะต้องลดปริมาณข้อมูลที่มีอยู่ในภาพเคลื่อนไหวให้ต่ำลงโดยให้มีคุณภาพในระดับยอมรับได้เพื่อให้มีความเป็นไปได้ทางเศรษฐกิจ และนับว่าเป็นโชคดีอย่างหนึ่งที่ภาพเคลื่อนไหวที่เราต้องการให้บริการในระบบโทรศัพท์นั้นมักจะเป็นภาพเคลื่อนไหวที่มีการเปลี่ยนแปลงของภาพไม่รวดเร็วนัก ยกตัวอย่างเช่นภาพของห้องประชุม การเคลื่อนไหวนั้นจะเกิดจากการเคลื่อนไหวตัวของสมาชิกในห้องประชุมซึ่งจัดว่าไม่รวดเร็วและไม่มากนัก สำหรับฉากหน้าและฉากหลังทั้งหลายจะอยู่นิ่ง ลักษณะดังกล่าวนี้ทำให้สัญญาณภาพในแต่ละเฟรมมีความสัมพันธ์ระหว่างกันสูงมาก หรือกล่าวได้ว่ามีความซ้ำซ้อนของข่าวสาร (message redundancy) สูงมาก นั่นหมายถึงถ้าเราสามารถขจัดความซ้ำซ้อนของข่าวสารออกไปได้ก็จะลดปริมาณข้อมูลลงได้ เกี่ยวกับวิธีการลดปริมาณข่าวสารของภาพเคลื่อนไหวนี้จะกล่าวในหัวข้อต่อไป

2.2 การลดปริมาณข้อมูลของภาพเคลื่อนไหว

จากที่อธิบายไว้ในหัวข้อ 2.1 จะสรุปได้ว่าสาเหตุที่ทำให้เราต้องส่งสัญญาณโทรทัศน์ด้วยอัตราข้อมูลที่สูงมากนั้นเป็นเพราะว่าสัญญาณเบสแบนด์ของภาพเคลื่อนไหวมี

ความถี่สูงกว่า 4 MHz และการส่งแบบดิจิตอลแบบ PCM ทำให้ต้องการอัตราข้อมูลสูงขึ้น อย่างไรก็ตามถ้าภาพเคลื่อนไหวที่ต้องการส่งมีความซ้ำซ้อนของข่าวสารอยู่เมื่อเราสามารถขจัดความซ้ำซ้อนออกไปได้ก็จะลดปริมาณข้อมูลลงได้ ความซ้ำซ้อนของข่าวสารในภาพเคลื่อนไหวนั้นเกิดขึ้นจากคุณสมบัติของภาพเคลื่อนไหวหลัก ๆ 3 ประการคือ (1) ภาพในเฟรมที่อยู่ติดกันมีความสัมพันธ์กันสูง (2) สเปกตรัมของความถี่ต่ำมีความหนาแน่นมากกว่าสเปกตรัมของความถี่สูง (3) สัญญาณภาพในแต่ละเฟรมเมื่อแบ่งเป็นบล็อกย่อย ๆ ขนาด 8x8 จุด (pixel) หรือ 16x16 จุดจะมีจำนวนแบบของแพตเทินที่จำกัด

คุณสมบัติประการแรกนั้นเกิดขึ้นสำหรับภาพเคลื่อนไหวทุกชนิดและจะเกิดขึ้นมากสำหรับภาพของห้องประชุมที่มีฉากหน้าและฉากหลังไม่เปลี่ยนแปลง ลักษณะดังกล่าวนี้ทำให้เราสามารถลดความซ้ำซ้อนของข่าวสารลงได้โดยนำสัญญาณที่อยู่ในเฟรมติดกันมาลบออกจากกัน แล้วจึงนำผลต่างที่ได้ไปเข้ารหัส PCM ต่อไป ซึ่งวิธีการนี้เรียกว่า DPCM (Differential PCM) การทำ DPCM สามารถลดปริมาณข้อมูลลงได้เมื่อเทียบกับการทำ PCM เพราะว่าผลต่างของสัญญาณดังกล่าวมีขนาดเล็กลงกว่าสัญญาณเดิม จึงสามารถใช้จำนวนบิตน้อยลงในการเข้ารหัสได้ และเพื่อให้การมอดูเลตมีประสิทธิภาพที่สูงขึ้นจะมีการทำ ADPCM (Adaptive DPCM) ซึ่งหมายถึงการทำ DPCM ที่เปลี่ยนแปลงให้เหมาะสมกับระดับสัญญาณที่เข้ามา ทำให้สามารถลดปริมาณข้อมูลที่ต้องส่งลงได้อีก

คุณสมบัติประการที่สองนั้นเป็นเพราะสเปกตรัมของความถี่ต่ำเกี่ยวข้องกับความสว่างของภาพในขณะที่สเปกตรัมของความถี่สูงเกี่ยวข้องกับรายละเอียดของภาพ ดังนั้นถ้าภาพเคลื่อนไหวนั้นมีรายละเอียดไม่สูงมาก หรือรายละเอียดด้านสูงถูกตัดออกไปบ้างภาพที่ได้ก็ยังมีคุณสมบัติดีพอที่สามารถใช้งานได้ การที่เราสามารถตัดรายละเอียดของภาพออกได้นั้นก็เท่ากับเป็นการลดความถี่สูงสุดของสัญญาณเบสแบนด์ลงไป ซึ่งหมายถึงจะสามารถทำการแปลงเป็นสัญญาณดิจิตอลด้วยอัตราข้อมูลต่ำลง ในกรณีของภาพของห้องประชุมนั้นรายละเอียดของภาพส่วนใหญ่จะอยู่นิ่ง เพราะฉะนั้นการทำ DPCM หรือ ADPCM ที่กล่าวไว้ข้างต้นจะเป็นการตัดสเปกตรัมของความถี่สูงไปได้ในตัว

คุณสมบัติประการที่สามนั้นเกิดขึ้นเนื่องจากเฟรมแต่ละเฟรมของภาพเคลื่อนไหวนั้นเมื่อนำมาแบ่งเป็นบล็อกเล็ก ๆ ขนาด 8x8 จุด หรือ 16x16 จุด เรามักจะพบว่าบล็อกของจุดภาพดังกล่าวนี้จะมีแพตเทินที่ซ้ำ ๆ กันเป็นจำนวนแบบที่จำกัด ดังนั้นถ้าเรา

เตรียมแพตเทินที่เกิดขึ้นเอาไว้ที่ต้นทางและปลายทางของช่องสัญญาณ แทนที่เราจะส่งบล็อกของข้อมูลนั้นไปตามปกติ เราเพียงแต่ส่งรหัสของบล็อกข้อมูลนั้นไปก็เพียงพอเพราะอุปกรณ์ที่ปลายทางจะสามารถดึงเอาแพตเทินของบล็อกข้อมูลนั้นออกมาแสดงได้นั่นคือเราจะสามารถลดปริมาณข้อมูลที่ต้องส่งลงไปได้ วิธีลดปริมาณข้อมูลแบบนี้เรียกว่า Vector quantization โดยเขียนย่อว่า VQ [3]

ในการลดปริมาณข้อมูลโดยใช้คุณสมบัติประการที่สองของภาพเคลื่อนไหวนั้นในทางปฏิบัติจะใช้วิธีที่เรียกว่าทรานส์ฟอร์มโค้ดดิง (transform coding) ซึ่งหมายถึงการนำเอาบล็อกของจุดภาพมาทรานส์ฟอร์มให้อยู่ในอาณาจักรของความถี่ แล้วพิจารณาให้น้ำหนักแก่สเปกตรัมที่เกิดขึ้น และทำการเข้ารหัสอย่างเหมาะสม วิธีทรานส์ฟอร์มที่นิยมใช้กันในเรื่องของภาพนั้นคือ วิธี 2D-WHT (Two dimension Walsh-Hadamard transform) และวิธี 2D-DCT (Two dimension discrete cosine transform) แต่เนื่องจากวิธี 2D-DCT มี Energy packing และ Decorrelation สูงที่สุด [4] ทำให้สามารถลดขนาดของข้อมูลได้มากและให้ภาพที่มีคุณภาพดี วิธี 2D-DCT จึงเป็นวิธีที่นิยมใช้กันมากกว่า

วิธีลดปริมาณข้อมูล 3 แบบหลัก ๆ ที่กล่าวมานี้จะมีที่ใช้ทั่วไป โดยอาจจะใช้วิธีใดวิธีหนึ่งแยกอิสระ หรือใช้วิธี 2 วิธีผสมกันแบบไฮบริดก็ได้ อย่างไรก็ตามแต่ละวิธีก็มีลักษณะเด่นและด้อยจึงต้องเลือกใช้ให้เหมาะสมกับงาน เมื่อเปรียบเทียบวิธีทั้งสามนี้ วิธี DPCM เป็นวิธีที่ค่อนข้างง่ายแต่ก็มีความสามารถจำกัด วิธี VQ เป็นวิธีที่เมื่อทำได้ดีแล้วจะได้ผลดีมาก แต่กรรมวิธีมักจะยุ่งยาก กล่าวคือการสร้าง code book จะต้องใช้เวลาในการคำนวณมาก และ Code book ชุดหนึ่งที่สร้างขึ้นมาจากภาพประเภทหนึ่ง อาจจะไม่เหมาะสำหรับภาพประเภทอื่น ได้ ซึ่งผลที่เกิดขึ้นก็คือคุณภาพของภาพที่ได้รับที่ปลายทางจะด้อยลงไป นอกจากนั้นอุปกรณ์ต้นทางจะต้องใช้เวลาในการคำนวณมากเช่นเดียวกัน สำหรับวิธีทรานส์ฟอร์มโค้ดดิงนั้นเป็นการคำนวณ 2D-DCT ซึ่งมีแอลกอริทึมมาตรฐานอยู่และใช้ได้สำหรับภาพเคลื่อนไหวทุกแบบ วิธีนี้จึงเป็นวิธีที่เหมาะสำหรับงานทั่วไป และการทำ DCT นั้นสามารถทำได้ด้วยซอฟท์แวร์ และฮาร์ดแวร์ที่ใช้ DSP chip รุ่นใหม่ ๆ ได้ (DSP ย่อจาก Digital signal processing)

ในวิทยานิพนธ์ฉบับนี้จะใช้วิธีลดปริมาณข้อมูลแบบ DPCM ร่วมกับแบบทรานส์ฟอร์มโค้ดดิงตามที่จะกล่าวในหัวข้อต่อไป

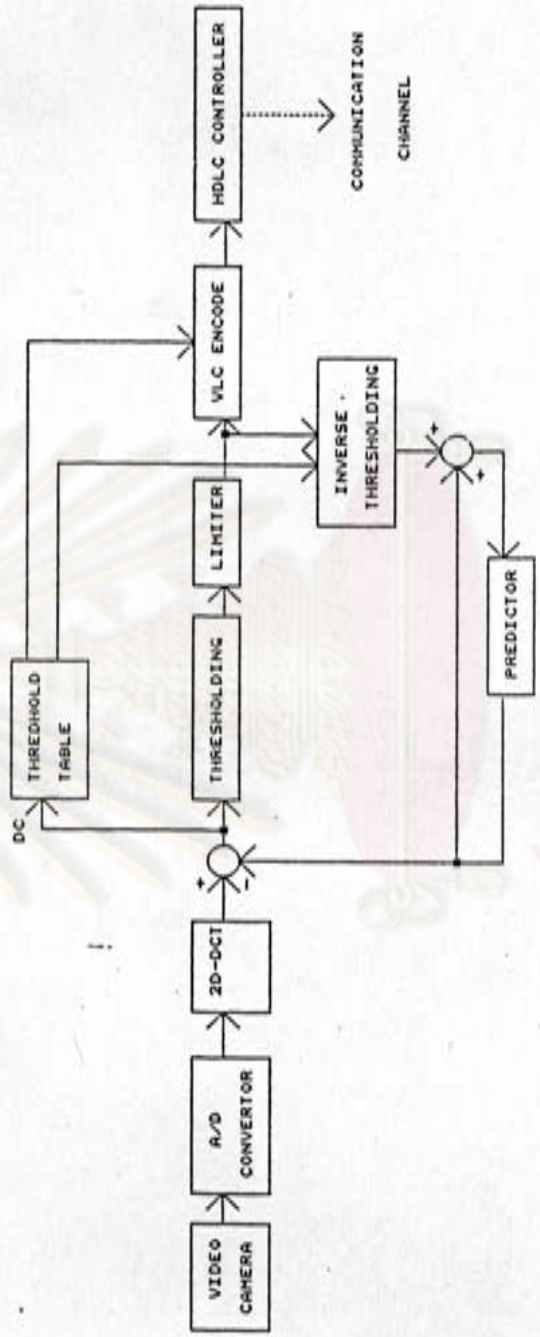
2.3 ส่วนประกอบและหลักการทำงานของระบบที่เสนอขึ้น

วิทยานิพนธ์ฉบับนี้จะเป็นการศึกษาการลดปริมาณข้อมูลของภาพเคลื่อนไหว วิธีที่เสนอใช้ในวิทยานิพนธ์นี้ได้แก่วิธี DPCM ผสมกับวิธีทรานส์ฟอร์มโค้ดดิง สาเหตุที่ต้องใช้สองวิธีประกอบกันนั้นเป็นเพราะต้องการลดปริมาณข้อมูลให้ได้มากที่สุด เพื่อให้เห็นตัวเลขชัดเจนขึ้นจะยกตัวอย่างที่เป็นจริงแสดงดังนี้คือ ในการส่งภาพขาวดำ 1 ภาพที่มีขนาด 128x128 จุด และแต่ละจุดมีระดับความสว่าง 256 ระดับ หรือ 8 บิต จะมีขนาดข้อมูลถึง 131,072 บิต ซึ่งปริมาณข้อมูลขนาดนี้ถ้าส่งผ่านช่องสัญญาณขนาด 64 kbps ภาพขาวดำ 1 ภาพนี้จะต้องใช้เวลาส่งมากกว่า 2 วินาที แต่การที่จะให้สายตาคนเราพอยอมรับได้ว่าเป็นภาพเคลื่อนไหวได้นั้นจะต้องมีการเปลี่ยนเฟรมของภาพไม่ต่ำกว่า 6 ภาพต่อวินาที นั่นคือจะต้องมีการลดปริมาณข้อมูลของภาพลงไม่ต่ำกว่า 12 เท่า ซึ่งจะทำได้ลำบากโดยใช้วิธีใดวิธีหนึ่งเพียงวิธีเดียว และเมื่อคำนวณค่าเฉลี่ยของจำนวนบิตต่อหนึ่งจุดภาพจะพบว่าเราจะต้องสามารถลดปริมาณข้อมูลจาก 8 บิตต่อจุดภาพเป็น 0.5 บิตต่อจุดภาพจึงสามารถส่งเป็นภาพเคลื่อนไหวในระดับยอมรับได้

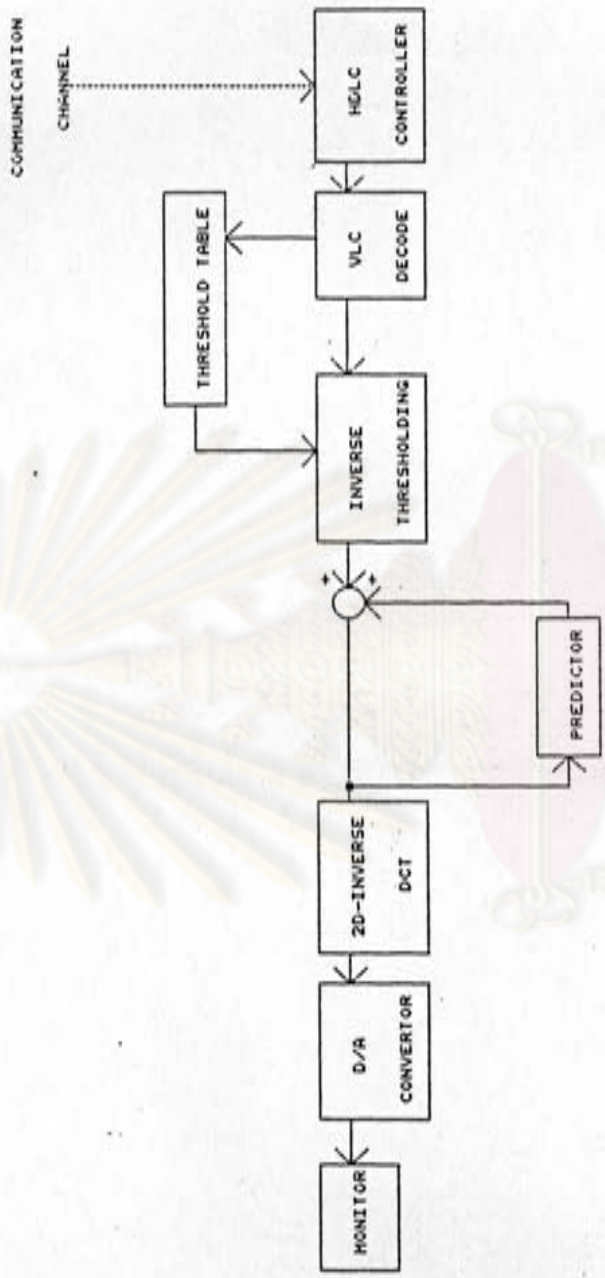
2.3.1 โครงสร้างของระบบ ระบบที่เสนอขึ้นมีโครงสร้างของภาคส่งดังที่แสดงไว้ในรูปที่ 2.1 และ ภาครับแสดงในรูปที่ 2.2

2.3.1.1 ภาคส่ง ในภาคส่งจะประกอบด้วยส่วนประกอบหลัก ๆ คือ A/D converter ซึ่งทำหน้าที่เปลี่ยนสัญญาณภาพจากกล้องวิดีโอในรูปสัญญาณแอนาลอกให้เป็นสัญญาณดิจิตอล 2D-DCT (2 Dimensions Discrete Cosine Transform) ซึ่งเป็นการดึงเอาสัญญาณภาพในรูปสัญญาณดิจิตอลมาทำ 2D-DCT โดยดึงเป็นบล็อกขนาด 8x8 จุด ส่วนต่อไปคือการทำ ADPCM แล้วจึงนำผลที่ได้ไปเข้ารหัสแบบ VLC (variable length code) ซึ่งจะสามารถลดปริมาณข้อมูลลงได้อีก

ในขั้นตอนของการทำ ADPCM นั้นเนื่องจากขั้นตอนนี้ตามหลังการทำ 2D-DCT ดังนั้นจึงเป็นการทำ ADPCM ในอาณาจักรของความถี่ ซึ่งจะให้ผลในการแสดงปริ



รูปที่ 2.1 โครงสร้างของภาคส่ง



รูปที่ 2.2 โครงสร้างของภาครับ

มาณข้อมูลเช่นเดียวกับการทำ ADPCM ในอาณาจักรของเวลา การทำเทรชโฮลดิง (thresholding) เป็นส่วนหนึ่งของกระบวนการทำ ADPCM ซึ่งทำให้มีการใช้จำนวนบิตที่มีประสิทธิภาพยิ่งขึ้น และข้อมูลส่วนนี้จะถูกเข้ารหัสและส่งไปยังภาครับด้วย

2.3.1.2 ภาครับ ในภาครับนั้นการทำงานจะย้อนรอยกระบวนการในภาคส่งกล่าวคือ จะประกอบด้วย VLC decoder และ Inverse threshoding และ Predictor ซึ่งจะทำการย้อนทาง ADPCM หลังจากนั้นจึงทำ Inverse 2D-DCT แล้วจึงผ่าน D/A convertor แปลงกลับไปเป็นสัญญาณแอนาลอกเพื่อแสดงที่จอภาพต่อไป

2.3.2 <u>หลักการทำงานของระบบ</u> การทำงานของระบบแบ่งได้เป็น 2 ส่วนใหญ่ๆ ดังนี้

2.3.2.1 ภาคส่ง ภาพจากกล้องวิดีโอจะถูกแปลงเป็นสัญญาณ ดิจิตอล โดยมีขนาด 128x128 จุดๆละ 6 บิตเก็บลงใน Vodeo Memory และจะถูก แบ่งเป็น block ขนาด 8x8 จุดเพื่อส่งไปทำ 2-D DCT ข้อมูลที่ผ่านการ Transform แล้วจะถูกส่งออกแบบ zigzag scan ตามรูปที่ 2.3 ซึ่งเป็นการเรียงข้อมูลตามลำดับความถี่ โดยข้อมูลที่อยู่มุมบนซ้าย จะเป็นค่า DC ของ block นั้น ตัวถัดไปตามแนว zigzag จะมีความถี่สูงขึ้นตามลำดับ โดยข้อมูลที่อยู่มุมขวาล่างคือส่วนประกอบของความถี่ที่สูงที่สุด การเรียงข้อมูลแบบนี้ทำให้สามารถเลือกส่งส่วนประกอบความถี่ต่ำซึ่งเป็นเค้าโครงของภาพไปก่อนได้แล้วจึงส่งส่วนประกอบความถี่สูงซึ่งเป็นรายละเอียดของภาพตามไป หรือจะตัดส่วนประกอบความถี่สูงทิ้งเพื่อลดข้อมูลก็ได้ถ้าคุณภาพของภาพไม่ด้อยลงมากนัก

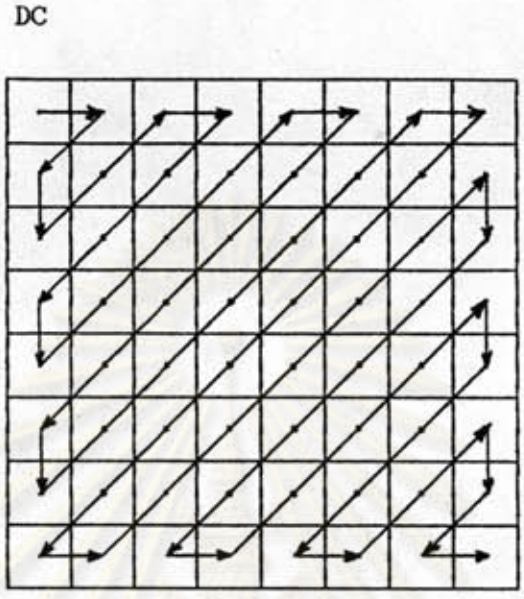
ข้อมูลที่ถูกส่งออกจะถูกนำมาลบกันจุดต่อจุดกับ Predictor ซึ่งเป็นข้อมูลของภาพที่แล้วใน Transform domain ผลต่างที่ได้จะผ่าน Thresholding โดยข้อมูลตัวแรกของแต่ละ block ซึ่งคือค่า DC ของ block นั้นจะเป็นตัวกำหนดค่า Threshold ที่จะนำมาทำ Thresholding ซึ่งมีหลักการดังนี้

ข้อมูลที่ผ่านการทำ 2-D DCT แล้วส่วนใหญ่จะมีค่าน้อยกว่า 10 [5] ดังนั้นถ้าเราเลือกค่า Threshold T มาทำการ encode โดย

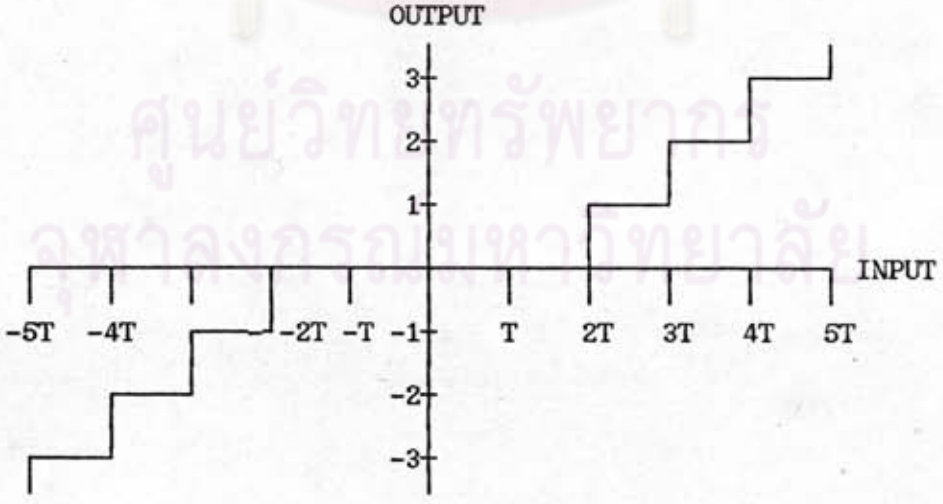
$$\text{OUTPUT} = (\text{INPUT} - T)/T \quad \text{เมื่อ INPUT} > T$$

$$= 0 \quad \text{เมื่อ INPUT} <= T$$

ซึ่งกรรมวิธีนี้จะมีผลตอบตามรูปที่ 2.4



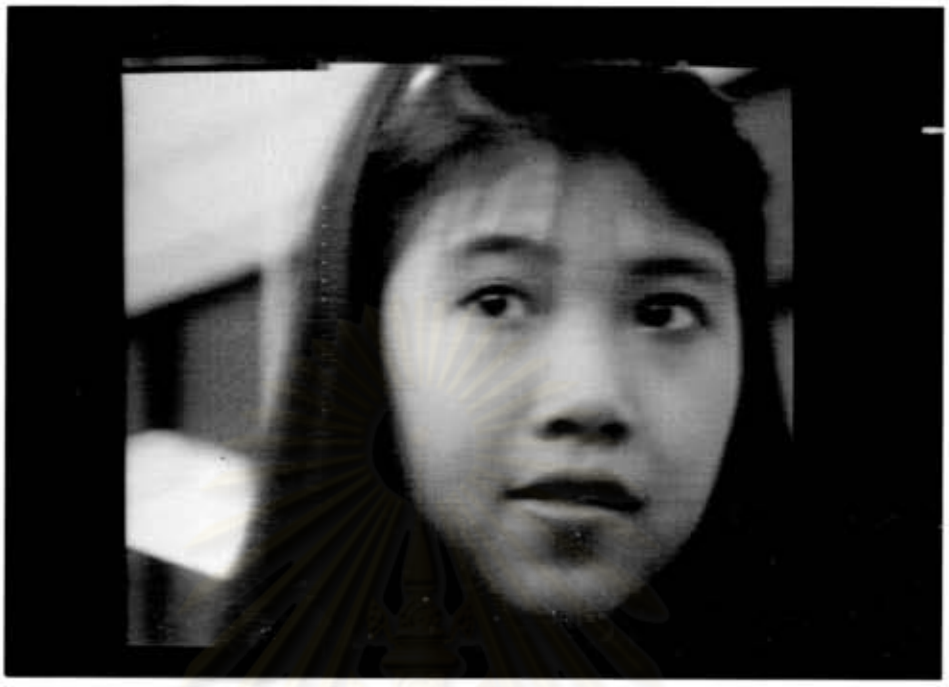
รูปที่ 2.3 การอ่านข้อมูลแบบ zigzag scan



รูปที่ 2.4 กราฟแสดงผลตอบของวิธี Thresholding

ถ้า Threshold มีค่ามากข้อมูลส่วนใหญ่จะมีค่าเป็น 0 ซึ่งเมื่อผ่านการเข้ารหัสแบบ Huffman แล้วข้อมูลที่ต้องส่งจะมีน้อยลง ถ้า Threshold มีค่าน้อยจะทำให้ข้อมูลที่ต้องส่งมีจำนวนมาก ดังนั้นค่า Threshold จึงสามารถใช้เป็นตัวกำหนดอัตราของข้อมูลได้ และจากการทดลองเก็บค่า DCT Coefficients ของภาพ L40.PIC และภาพ KUMIKO.PIC ดังรูปที่ 2.5 และ 2.6 ตามลำดับ แล้วนำมาพล็อตกราฟเพื่อดูการกระจายของ ค่า DCT Coefficients ตาม Threshold ค่าต่างๆ ดังรูปที่ 2.7 จะเห็นได้ว่า ถ้าเลือกค่า Threshold สูงกว่า 14 จำนวนข้อมูลซึ่งน้อยกว่าค่า Threshold จะมีถึง 90% ซึ่งเมื่อผ่านการทำ Thresholding แล้ว จะมีค่าเป็น 0 และจากการทดลองเปลี่ยนค่า Threshold แล้วดูคุณภาพของภาพและจำนวนข้อมูลพบว่า ค่า Threshold ที่เหมาะสมอยู่ระหว่าง 4 และ 16 โดยค่า Threshold เป็น 4 สามารถลดข้อมูลได้พอสมควรโดยดูจากกราฟในรูปที่ 2.7 จะเห็นว่า ถ้าใช้ค่า Threshold เป็น 4 ข้อมูลประมาณ 75% จะเป็น 0 และจากการทดลองจำลองแบบการทำงานด้วยโปรแกรม ดังแสดงในภาคผนวก ก พบว่า คุณภาพของภาพไม่ด้อยลงมากนัก และถ้าใช้ค่า Threshold เป็น 16 จะสามารถลดข้อมูลลงได้มากโดยข้อมูลมากกว่า 90% จะเป็น 0 แต่จะมีผลต่อคุณภาพของภาพมากคือภาพที่ได้จะมีลักษณะเป็นบล็อก จึงใช้ค่า Differential ของ DC มาเป็นตัวกำหนด Threshold T โดยที่ ถ้า DC มีค่ามากแสดงว่า block นั้นมีการเปลี่ยนแปลงจาก block เดียวกันของ frame ที่แล้วมากจึงตั้งค่า Threshold ให้มีค่ามากเพื่อให้สามารถส่งเค้าโครงส่วนใหญ่ได้ทัน แต่ถ้า DC มีค่าน้อยก็จะตั้งค่า Threshold ให้ต่ำเพื่อที่จะส่งส่วนที่เป็นรายละเอียดของภาพ ดังนั้นถ้าภาพเคลื่อนไหวมากก็จะเบลอแต่จะยังคงมีเค้าโครงส่วนใหญ่ของภาพอยู่เมื่อภาพหยุดนิ่งหรือเคลื่อนไหวน้อยก็จะค่อยๆมีรายละเอียดของภาพปรากฏชัดขึ้น โดยอัตราการรับส่งจะค่อนข้างคงที่ ดังนั้นในวิทยานิพนธ์ฉบับนี้จึงใช้ค่า Threshold 2 ค่าคือ 4 และ 16 โดยถ้าค่า DC มีค่าตั้งแต่ 32 ขึ้นไปก็จะใช้ค่า T เป็น 16 แต่ถ้าน้อยกว่า 32 ก็จะใช้ค่า T เป็น 4 และค่า T นี้จะถูกเขียนลงใน Threshold Table เพื่อนำไปเข้ารหัสและส่งไปยังภาครับด้วย

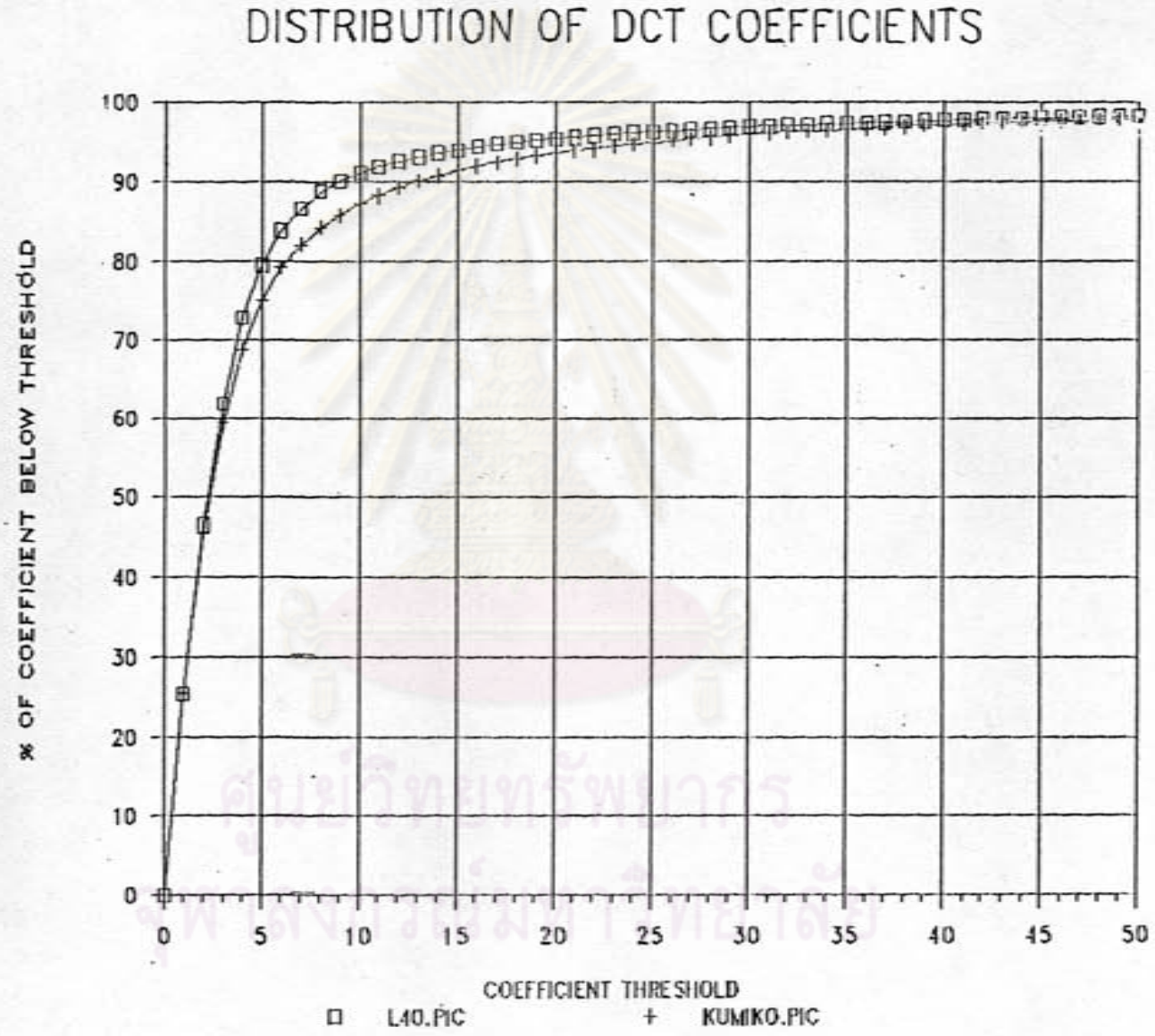
หลังจากการทำ Thresholding แล้วข้อมูลจะถูกจำกัดด้วย Limiter โดยจะจำกัดให้อยู่ในช่วง -32 ถึง 31 เพื่อไม่ให้จำนวนบิตของข้อมูลมากเกินไป ข้อมูลที่ผ่าน Limiter แล้วจะถูกแยกเป็น 2 ทาง ทางหนึ่งจะผ่าน Inverse



รูปที่ 2.5 ภาพ L40.PIC

รูปที่ 2.6 ภาพ KUMIKO.PIC

018077



รูปที่ 2.7 กราฟแจกแจงของค่า DCT Coefficients ตาม Threshold ต่างๆ

Thresholding เพื่อไปรวมกับค่าใน Predictor เพื่อเป็น Predictor ของ frame ต่อมา อีกทางหนึ่งจะผ่านไปทำ VLC Encode ซึ่งเป็นการเข้ารหัสแบบ Huffman เพื่อลดจำนวนของข้อมูล โดยข้อมูลจะแยกเป็น amplitude กับ run-length ของ 0 ดังตารางที่ 2.1 และ 2.2 ข้อมูลที่เข้ารหัสแล้วจะถูกส่งออกด้วยอัตรา 64 kb/s ผ่านช่องทางสื่อสารไปยังเครื่องรับ

2.3.2.2 ภาครับ ข้อมูลที่เข้ามาที่ภาครับจะผ่าน VLC Decoder ทำการถอดรหัส Huffman แยกค่า Threshold(T) ออกเขียนลงใน Threshold Table แล้วจึงทำ Inverse Thresholding ย้อนรอยกรรมวิธีทางภาคส่ง แล้วนำข้อมูลที่ได้ไปบวกกับข้อมูลใน Predictor ซึ่งเป็นการย้อนรอย DPCM ข้อมูลที่ได้ซึ่งเป็นการเรียงลำดับแบบ zigzag จะถูกเรียงกลับเป็น block 8x8 แล้วจึงส่งไปทำ 2-D DCT^{-1} และผ่าน D/A Convertorเพื่อแปลงกลับเป็นภาพแสดงยังจอภาพ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ขนาดของข้อมูล	จำนวนบิตขาออก	รหัสขาออก
1	1	1
2	3	001
3	4	0111
4	5	0001
5	5	01101
6	6	011001
7	7	0000001
8	7	0110001
9	8	00000000
10	8	01100000
11	8	00000001
12	8	01100001
ตั้งแต่ 13	6+6	000001+ข้อมูล
EOB	4	0001
รหัสนำหน้าข้อมูลซ้ำ	3	010

ตารางที่ 2.1 การเข้ารหัสแบบ Huffman ของ Amplitude

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

จำนวนข้อมูลที่มีค่าเป็น 0 ที่ติดกัน	จำนวนบิตขาออก	รหัสขาออก
1	2	11
2	3	101
3	3	011
4	4	0101
5	4	0011
6	5	01000
7	5	10010
8	5	01001
9	5	10001
10	5	10011
11	6	001000
12	6	100000
13	6	001010
14	6	001001
15	6	100001
16	6	000011
17	6	001011
18	7	0000000
19	7	0000100
20	7	0000010
21	7	0001110

ตารางที่ 2.2 การเข้ารหัสแบบ Huffman ของ ข้อมูลที่เป็นศูนย์ที่ติดกัน

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

จำนวนข้อมูลที่มีค่าเป็น 0 ที่ติดกัน	จำนวนบิตขาออก	รหัสขาออก
22	7	0000001
23	7	0000101
24	7	0000011
25	7	0001111
26	8	00011000
27	8	00011010
28	8	00011001
29	8	00011011
30	5+6	00010+จำนวน

ตารางที่ 2.2 (ต่อ)การเข้ารหัสแบบ Huffman ของ ข้อมูลที่เป็นศูนย์ที่ติดกัน

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย