บทที่ 2
การรับส่งภาพนิ่งแบบโปรเกรสซีฟ

ในบทนี้จะกล่าวถึงจุดมุ่งหมาย ความสำคัญ และ หลักการเบื้องต้นของการส่งภาพนิ่งแบบโปรเกรสซีฟ เพื่อเป็นแนวทางสำหรับการศึกษา และ พัฒนา เพื่อนำไปประยุกต์ใช้อย่างเหมาะสมต่อไป นอกจากนั้นก็จะกล่าวถึงโครงสร้างแบบมีลำดับชั้น (Hierachical Structure) ซึ่งจะเป็นวิธีที่ใช้ในการสร้างระบบรับส่งภาพนิ่งในวิทยานิพนธ์ฉบับนี้ และ มาตรฐานของกลุ่ม JPEG (Joint Photographic Experts Group) ซึ่งเป็นกลุ่มทำงานที่จัดตั้งขึ้นมาเพื่อสร้างมาตรฐานในการรับส่งภาพนิ่ง

2.1 ลักษณะของการรับส่งภาพแบบโปรเกรสซีฟ

การรับส่งภาพแบบโปรเกรสซีฟ อาจจะกล่าวได้ว่าเป็น การนำเอาข้อมูลข่าวสารที่มีอยู่ในภาพหนึ่ง ๆ มาจัดเรียงลำดับในการส่งเสียใหม่ เพื่อเพิ่มประสิทธิภาพในกระบวนการส่งข้อมูลไปตามช่องสัญญาณที่มีแบนด์วิดท์จำกัด การเรียงลำดับข้อมูลจะทำให้ลักษณะของภาพที่รับได้ปรากฏเป็นภาพคร่าว ๆ ของภาพต้นแบบที่จะค่อย ๆ ดีขึ้นตามลำดับเวลา

การรับส่งภาพแบบนี้ จะใช้ความสามารถของมนุษย์ในการเข้าใจภาพ (image understanding) เป็นหลัก วิธีการเรียงลำดับข่าวสารแบบนี้จะทำให้ผู้ใช้งานที่อยู่ไกลออกไปซึ่งมักจะเป็นมนุษย์ ประหยัดเวลาในการค้นหาเป็นอย่างมาก ผู้ใช้อาจจะรู้ว่าภาพที่ต้องการจะมีลักษณะเป็นอย่างไร ทำให้สามารถโต้ตอบได้อย่างเหมาะสม (เช่น ต้องการ/ไม่ต้องการ) โดยต้องการข้อมูลก่อนการตัดสินใจ น้อยกว่าข้อมูลของภาพต้นแบบมาก ๆ

2.2 ข้อดีและข้อเสียในการส่งภาพแบบโปรเกรสซีฟ

การส่งภาพแบบโปรเกรสซีฟมีข้อดีกว่าการส่งแบบ slow scan ที่เห็นได้ชัดเจนคือ

1. ในแต่ละขั้นตอน ภาพที่ปรากฏขึ้นทางด้านรับจะมีขนาดเท่ากับภาพต้นแบบตลอดเวลา แทนที่จะเป็นเพียงบางส่วนของภาพ ทำให้ผู้ใช้ได้รับความรู้สึกที่ดีกว่า และ ยังทำให้เห็นลักษณะของภาพอย่างหยาบ ๆ เพื่อเป็นข้อมูลในการตัดสินใจเลือกภาพที่ต้องการ

2. ในบางกรณีที่เราไม่ต้องการภาพที่มีความละเอียดสูงที่สุด ก็อาจจะสั่งให้หยุดส่งได้ทันที แล้วนำภาพที่รับได้ในขณะนั้นไปใช้ต่อไป ทำให้สามารถลดเวลาในการส่งลงได้อีก ซึ่งเป็นข้อดีที่การส่งแบบ Slow scan ไม่สามารถทำได้ เพราะการนำไปใช้งานจะต้องรอให้ภาพปรากฏเต็มจอภาพเสียก่อน

ส่วนข้อเสียที่อาจเกิดขึ้นกับการส่งแบบโปรเกรสซีฟบางวิธี ก็คือ ในการเรียงลำดับของการส่งข้อมูลใหม่นั้น จำเป็นต้องมีส่วนของข้อมูลเพิ่มเติม (Overhead) ส่งไปนอกเหนือจากข้อมูลที่มีอยู่เดิม เพื่อควบคุมการเรียงลำดับข้อมูล ทำให้ต้องเสียเวลาในการส่งส่วนนี้เพิ่มตามไปด้วย ดังนั้นเวลาที่ใช้ตั้งแต่เริ่มต้นส่งจนสิ้นสุดการส่งก็อาจจะนานกว่าการส่งในแบบ slow scan ได้ ถ้าข้อมูลรวมทั้งหมดมากกว่าข้อมูลของภาพต้นแบบ

2.3 หลักการเบื้องต้น

การส่งแบบโปรเกรสซีฟจะส่งข้อมูลให้กับด้านรับไปเรื่อย ๆ จนกว่าจะเพียงพอให้ผู้ใช้ตัดสินใจได้ว่าต้องการภาพดังกล่าว หรือไม่ แต่จะต้องพยายามให้ผู้ใช้ใช้เวลาตัดสินใจให้เร็วที่สุด การสร้างภาพในระยะแรก ๆ จึงต้องอาศัยข้อมูลที่มีจำนวนบิตน้อยมาก ๆ ด้วยจำนวนบิตที่น้อยนี้เอง ลักษณะของภาพที่ปรากฏทางด้านรับในตอนแรก จะมีลักษณะคร่าว ๆ ไม่ชัดเจน รายละเอียดน้อย แต่จะต้องมีขนาดเท่ากับภาพต้นแบบ เพื่อจะให้ความรู้สึกที่ดีต่อผู้ใช้ดังที่กล่าวมาแล้ว เมื่อเวลาผ่านไปด้านรับจะได้ข้อมูลเพิ่มขึ้น ข้อมูลที่ได้รับเพิ่มจะถูกนำไปรวมกับข้อมูลเดิม เพื่อปรับปรุงให้ภาพมีรายละเอียดเพิ่มขึ้นเรื่อย ๆ ตามลำดับ ในการส่งภาพอาจจะมีการแบ่งเป็นขั้นตอนตามความละเอียดของภาพที่รับได้ โดยวัดเป็นจำนวนบิตต่อจุดที่เทียบกับภาพต้นแบบ โดยในขั้นสุดท้ายควรมีรายละเอียดที่แตกต่างจากภาพต้นแบบให้น้อยที่สุด

สำหรับการประยุกต์บางอย่างที่ต้องการรายละเอียดของภาพสูง ภาพที่ต้องการที่ด้านรับจะต้องไม่มีความเพี้ยน (distortion) จากภาพต้นแบบเลย เช่น ในระบบรักษาความปลอดภัย หรือ ภาพเอกซ์เรย์ทางการแพทย์ เป็นต้น ในกรณีนี้ก็อาจจะต้องการการส่งในแบบที่เรียกว่า lossless transmission คือ การส่งภาพในลักษณะที่ภาพด้านรับจะต้องไม่มีความเพี้ยนใด ๆ เกิดขึ้นเมื่อเทียบกับภาพต้นแบบ การส่งในลักษณะนี้มักจะใช้เสริมกับการส่งแบบโปรเกรสซีฟถ้าต้องการ โดยจะส่งเป็นขั้นสุดท้ายซึ่งอาจจะเรียกว่าขั้นของ Reversible scan เพื่อให้ด้านรับสามารถสร้างภาพที่เหมือนกับภาพต้นแบบได้อย่างสมบูรณ์

วิธีการส่งภาพแบบโปรเกรสซีฟ อาจแบ่งได้เป็น 2 วิธีใหญ่ ๆ คือ

1. วิธีทาง Spatial Domain คือ การเรียงข้อมูล หรือประมวลผลภาพโดยตรง เช่น DPCM (Differential Pulse Code Modulation), DM (Delta Modulation), VQ (Vector Quantization) เป็นต้น

2. วิธีทาง Spectral Domain คือ วิธีที่มีการทรานส์ฟอร์มข้อมูลภาพ โดยใช้ออร์โทโกนัลทรานส์ฟอร์มแบบใดแบบหนึ่ง แล้วนำไปประมวลผลใน Spectral Domain ต่อไป เช่น ฟูเรียร์ทรานส์ฟอร์ม (Fourier transform), โคซายน์ทรานส์ฟอร์ม (Cosine transform) เป็นต้น

วิธีทาง Spectral Domain เป็นวิธีที่นิยมใช้กันมากกว่าในการรับส่งภาพแบบโปรเกรสซีฟ เพราะใน Spectral Domain จะมีการอัดแน่นของพลังงาน (Energy Packing) ของข้อมูลในระดับความถี่ต่ำ ทำให้สามารถลดความซ้ำซ้อน (Redundancy) ของข้อมูลได้ง่าย และ การส่งในลักษณะแบบโปรเกรสซีฟก็ทำได้ง่ายเช่นเดียวกัน นอกจากนี้ยังสามารถนำเอาเทคนิคแบบ adaptive มาใช้ประกอบเพื่อเพิ่มประสิทธิภาพได้อีกด้วย

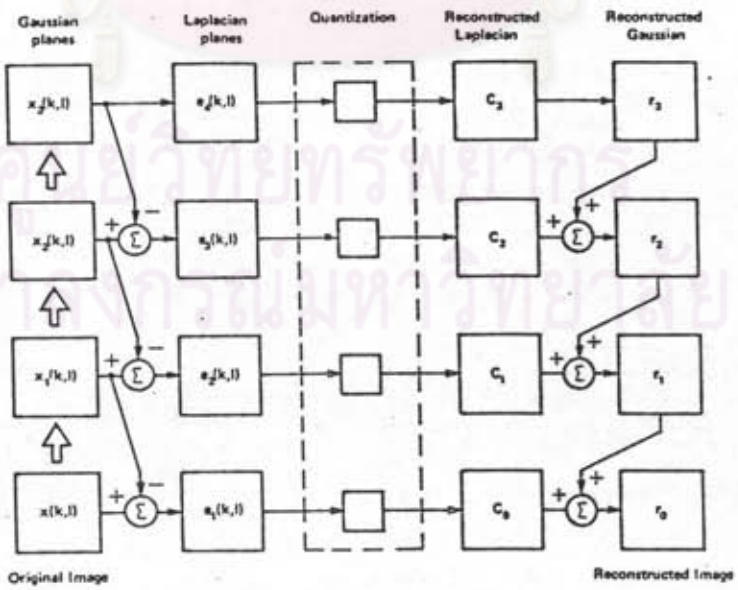
อย่างไรก็ตามการส่งภาพแบบโปรเกรสซีฟวิธีต่าง ๆ จะต้องยึดหลักดังต่อไปนี้

1. ในระหว่างการส่งข้อมูลในแต่ละขั้น จะต้องระวังไม่ให้มีการซ้ำกันของข่าวสาร หรือให้มีน้อยที่สุด เพราะจะทำให้เสียเวลาในการส่งเพิ่มขึ้น

2. ความเพี้ยนของภาพที่เกิดขึ้นในขั้นสุดท้ายจะต้องมีน้อยที่สุด หรือ ถ้าเป็นไปได้จะต้องไม่มีเลย

ในอดีตได้มีผู้ทดลอง และเสนอวิธีการรับส่งภาพแบบโปรเกรสซีฟอยู่หลายวิธีด้วยกัน K.Takaikawa [2] ได้เสนอวิธีการแยกสัมประสิทธิ์ใน Spectral Domain ออกเป็น Sparse matrix ซึ่งจะมีมิติที่แตกต่างกันไล่จากเล็กไปหาใหญ่ เมื่อเราทำการส่งสัมประสิทธิ์ที่มีความถี่ต่ำไปก่อนบางค่า สัมประสิทธิ์เหล่านี้จะถูกทรานส์ฟอร์มกลับ ซึ่งประมวลผลได้เร็ว ต่อมาเมื่อได้รับสัมประสิทธิ์เพิ่มขึ้นมิติที่จะถูกทรานส์ฟอร์มกลับก็จะมีขนาดใหญ่ขึ้นด้วย เมื่อนำผลที่ได้จากการทรานส์ฟอร์มกลับในแต่ละขั้นมารวมกัน ก็จะได้ภาพที่มีรายละเอียดชัดขึ้นเรื่อย ๆ อย่างไรก็ตามวิธีนี้จะไม่เหมาะกับภาพที่มีรายละเอียดมาก เพราะสัมประสิทธิ์ส่วนใหญ่จะมีความถี่สูง ในขณะที่ลำดับการส่งสัมประสิทธิ์จะต้องไล่จากความถี่ต่ำก่อน โดยที่ไม่อาจเปลี่ยนแปลงได้ ทำให้สร้างภาพได้ช้า Chen & Smith [3] ได้ใช้วิธี adaptive ใน spectral domain โดยการแบ่งภาพออกเป็นบล็อกย่อย ๆ ขนาดเท่า ๆ กัน นำไปทรานส์ฟอร์มโดยใช้ Discrete

Cosine Transform (DCT) แล้วจำแนกบล็อกย่อยเหล่านั้นออกเป็นชั้นต่าง ๆ กัน ตามระดับพลังงาน (energy) ในแต่ละบล็อก โดยระดับพลังงานแต่ละชั้นก็จะมีการจัดสรรจำนวนบิต (bit allocation) ต่างกันไป ชั้นที่มีพลังงานมากก็จะถูกกำหนดให้มีจำนวนบิตสูง วิธีนี้จะทำให้สามารถลดข้อมูลได้มาก ทำให้ได้อัตราข้อมูลต่ำ แต่ก็ไม่เหมาะที่จะนำมาสร้างระบบรับส่งภาพนิ่งแบบโปรเกรสซีฟ เพราะต้องรอรับสัมประสิทธิ์ให้หมดบล็อกเสียก่อน ทำให้เสียเวลารอนานก่อนที่จะเริ่มประมวลผลได้ Burt & Adelson [4] ได้ศึกษาการส่งแบบพีระมิด (pyramidal coding) ซึ่งก็คือการแยกส่งภาพตามช่วงของความถี่ในทาง spectral domain นั่นเอง โดยเริ่มส่งจากภาพที่มีความถี่ต่ำไปก่อน แล้วจึงตามด้วยภาพที่มีแถบความถี่สูงขึ้นตามลำดับ แต่ละช่วงความถี่จะส่งโดยใช้วิธี DPCM หรือ วิธีอื่น ๆ แล้วแต่ความเหมาะสม แผนภูมิแสดงวิธีการส่งแบบนี้แสดงไว้ดังรูปที่ 2.1 การส่งโดยการแยกภาพตามแถบความถี่ได้ถูกเสนออีกในลักษณะของ subband coding โดย Wood & Oneal [5] และ Hofmann & Troxel [6] ในเวลาต่อมา Lohscheller [7], Chitraprasert & Rao [8] ได้นำเอาผลตอบสนองต่อการมองเห็นของตามนุษย์ (human visual response) มาใช้ร่วมกับวิธีทาง spectral domain กล่าวคือ ตาของมนุษย์จะมีผลตอบสนองต่อความถี่ต่าง ๆ ไม่เท่ากัน โดยสร้างเป็นตัวถ่วงน้ำหนัก (weighting) คูณเข้ากับสัมประสิทธิ์ใน Spectral Domain ก่อนที่จะนำไปประมวลผลต่อไป วิธีนี้จะให้ภาพที่มีลักษณะดีมากที่อัตราของข้อมูลต่ำมาก ๆ และ ยังทำการส่งแบบโปรเกรสซีฟได้ไม่ยากอีกด้วย



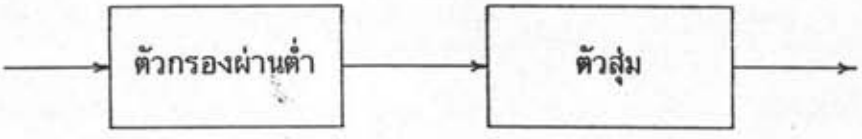
รูปที่ 2.1 แผนภูมิแสดงขั้นตอนในการเข้ารหัส และ ถอดรหัสแบบพีระมิด

ในปี 1986 CCITT และ ISO ได้ร่วมมือกันจัดตั้งกลุ่ม JPEG ขึ้น เพื่อหาแนวทางในการกำหนดการรับส่งภาพนิ่งให้เป็นมาตรฐาน JPEG ได้ทำการทดสอบวิธีการส่งหลายวิธี [9] และ ได้เสนอให้ใช้โครงสร้างแบบมีลำดับชั้นในการรับส่งแบบโปรเกรสซีฟดังจะกล่าวในหัวข้อถัดไป

2.4 โครงสร้างแบบมีลำดับชั้น (Hierachical Structure)

วิธีการส่งแบบโปรเกรสซีฟที่เสนอโดยกลุ่ม JPEG [10] นี้ จะประกอบไปด้วยกระบวนการ 2 อย่างคือ การใช้โครงสร้างแบบมีลำดับชั้น (Hierachical Structure) และ การเลือกส่งสัมประสิทธิ์ (Spectral Selection) โดยอาศัยกระบวนการทั้งสองอย่างนี้ เราสามารถกำหนดให้อัตราข้อมูลมีค่าเท่ากัน หรือ ใกล้เคียงกัน เพื่อจำกัดเวลาในการส่งภาพต่าง ๆ ในแต่ละขั้นให้เท่ากัน การส่งแบบนี้จะถูกกำหนดให้มีทั้งสิ้น 4 ขั้น โดยมีอัตราข้อมูลในแต่ละขั้นเป็น 0.08, 0.25, 0.75 และ 2.25 บิตต่อจุด จากภาพต้นแบบที่มีอัตราข้อมูลเป็น 8 บิตต่อจุด สังเกตว่าอัตราข้อมูลที่ได้ในขั้นแรกจะมีค่าน้อยมากเมื่อเทียบกับจำนวนบิตที่ต้องการสำหรับภาพต้นแบบ

ในขั้นแรก ๆ ที่มีอัตราข้อมูลต่ำมาก พบว่าถ้าใช้วิธีการควอนไตซ์หลังจากการทรานส์ฟอร์มอย่างธรรมดาทั่วไป ภาพที่ได้จากการประมวลผลกลับ จะมีคุณภาพต่ำกว่าภาพต้นแบบมาก เพราะผลของการแบ่งบล็อก (Block effect) ที่เกิดขึ้นเนื่องจากความละเอียดของสัมประสิทธิในแต่ละบล็อกที่มีค่าต่ำมาก วิธีการที่ให้ผลดีกว่าก็คือ การกรองผ่านต่ำภาพต้นแบบแล้วสุ่มให้มีจำนวนจุดลดน้อยลงดังรูปที่ 2.2 ก่อนที่จะไปทำการทรานส์ฟอร์ม และ ควอนไตซ์ ต่อไป ในกรณีนี้ภาพที่ได้หลังจากทำการประมวลผลกลับแล้ว จะพบว่า ความเพี้ยนของภาพที่เกิดขึ้นที่บริเวณขอบบล็อกเนื่องจากการประมาณค่าในช่วง (interpolate) จะน้อยกว่าในกรณีแรก เพราะว่าความละเอียดของสัมประสิทธิ์ตำแหน่งต่าง ๆ ในแต่ละบล็อกจะสูงกว่าในวิธีแรก



รูปที่ 2.2 การกรองผ่านต่ำ และ การสุ่มเพื่อลดจำนวนจุด

ตามหลักการส่งแบบโปรเกรสซีฟในหัวข้อ 2.3 เมื่อส่งข้อมูลของภาพในขั้นแรกไปแล้ว ก็ไม่จำเป็นต้องส่งข้อมูลในส่วนนั้นไปอีก ซึ่งทำได้โดยสร้างตัวทำนายที่จะทำนายภาพที่ปรากฏขึ้นทางด้านรับ แล้วนำภาพในตัวทำนายนั้นไปลบกับภาพต้นแบบที่ถูกสุ่มในขั้นต่อมา ผลต่างที่ได้จะเป็นภาพซึ่งจะถูกนำไปประมวลผลในลักษณะแบบเดียวกับขั้นแรก การสุ่มเพื่อลดจำนวนจุด และ การใช้ตัวทำนายนี้ จึงทำให้เกิดโครงสร้างแบบมีลำดับชั้นขึ้น

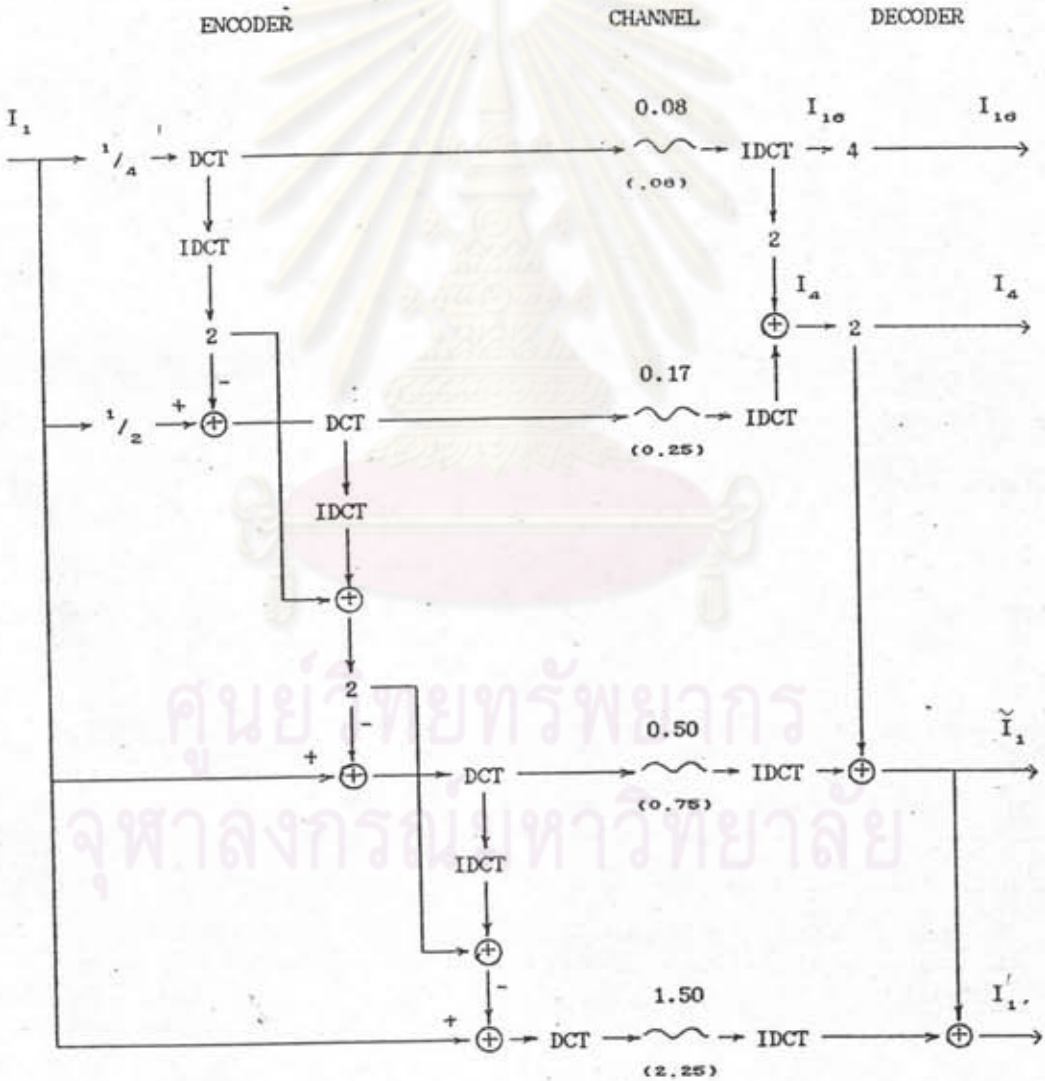
วิธีการดังกล่าวจึงคล้ายคลึงกับวิธีการส่งแบบพีระมิด (Pyramid Coding) แต่มีข้อดีกว่าตรงที่ได้มีการนำเอาภาพที่ได้จากตัวทำนายทางด้านส่ง ซึ่งเป็นภาพที่ด้านรับควรจะรับได้มาลบออก ทำให้ความผิดพลาดที่เกิดขึ้นจากการประมวลผลทั้งหลายจะถูกแก้ไขไปได้ในขั้นตอนถัดมา แต่อย่างไรก็ตาม ความเพี้ยนของภาพในขั้นตอนสุดท้ายก็ยังคงมีอยู่ และ ไม่อาจกำจัดให้หมดไปได้ ตราบใดที่ยังคงมีการควอนไตซ์อยู่ ในกรณีนี้ก็ต้องการการส่งอีกหนึ่งขั้นดังที่เรียกว่า Reversible scan ซึ่งเป็นการนำเอาความผิดพลาดที่ยังคงเหลืออยู่ในขั้นสุดท้าย ส่งไปตรง ๆ โดยใช้วิธีทาง Spatial Domain เช่น DPCM เป็นต้น

โดยปกติ รูปภาพที่ได้จากการทรานส์ฟอร์ม และ ควอนไตซ์นั้น ยังมีอัตราข้อมูลสูงอยู่มาก ในกรณีนี้การเข้ารหัสจึงเป็นเรื่องที่จำเป็น เพื่อลดอัตราข้อมูลลงมาให้อยู่ในระดับที่น้อยที่สุด การเข้ารหัสที่นิยมใช้มักจะเป็นการเข้ารหัสประเภทเอนโทรปี (entropy coding) เช่น Huffman เป็นต้น การส่งแบบมีลำดับชั้นจะมีกรรมวิธีการส่งดังรูปที่ 2.3

ในขั้นแรก ภาพต้นแบบจะถูกกรองและสุ่ม 1/4 เท่าในแต่ละแกน ทำให้ภาพมีขนาดเป็นเพียง 1/16 เท่าของภาพต้นแบบ แล้วนำภาพดังกล่าวไปทำการประมวลผลโดยวิธี ADCT (Adaptive Discrete Cosine Transform) ซึ่งเป็นการควอนไตซ์สัมประสิทธิ์ที่ได้จากการทรานส์ฟอร์ม DCT ด้วยเมตริกซ์ค่าขีดเริ่มของการมองเห็น (Visibility Threshold Matrix) ซึ่งก็เหมือนกับการเลือกสัมประสิทธิ์ที่มีนัยสำคัญจริง ๆ ออกมาก่อนนั่นเอง หลังจากนั้นก็จะนำสัมประสิทธิ์ที่เหลืออยู่ไปเข้ารหัสแบบ Hufmann ให้มีอัตราข้อมูลประมาณ 0.08 บิตต่อจุดตามที่ JPEG กำหนด แล้วส่งไปยังด้านรับ ทางด้านรับก็จะนำข้อมูลเหล่านี้ไปถอดรหัสแล้วทรานส์ฟอร์มกลับพร้อมกับประมาณค่า (interpolate) ขึ้นให้ได้ขนาดเท่ากับภาพต้นแบบ ขณะเดียวกันทางด้านส่งจะทำการประมวลผลเช่นเดียวกับด้านรับ แต่จะทำการประมาณค่าให้ภาพที่ได้มีขนาดเป็น 2 เท่าก่อนที่นำไปเก็บไว้ในตัวทำนายต่อไป

ในขั้นต่อมา ภาพต้นแบบจะถูกสุ่มและกรอง 1/2 เท่าในแต่ละแกน ซึ่งจะได้ขนาดภาพเท่ากับภาพที่อยู่ในตัวทำนายที่ได้จากขั้นแรก นำภาพทั้งสองมาลบกันจะได้ภาพแสดงความผิดพลาด (error) ซึ่งจะนำไปทำการทรานส์ฟอร์มด้วย DCT และ เข้ารหัสเช่นเดียวกับขั้นแรก ให้มีอัตราข้อมูลเป็น 0.17 บิตต่อจุดส่งไปยังด้านรับ ทำให้อัตราข้อมูลรวมในตอนนี้เป็น 0.25

บิตต่อจุด ด้านรับก็จะทำการประมวลผลเช่นเดียวกับขั้นแรก แล้วนำผลที่ได้ไปบวกกับภาพเดิม ทำให้ได้ภาพที่มีรายละเอียดดีขึ้น ในทำนองเดียวกับขั้นแรก ทางด้านส่งก็จะทำการประมวลผลกลับแล้วบวกเข้ากับตัวทำนายเดิม ประมาณค่าภาพที่ได้จากผลบวกให้มีขนาดเป็น 2 เท่า ก่อนที่จะเก็บไว้ในตัว ทำนายเพื่อใช้ในขั้นถัดไป ขนาดของภาพในตัวทำนายตอนนี้จึงเท่ากับขนาดของภาพต้นแบบ

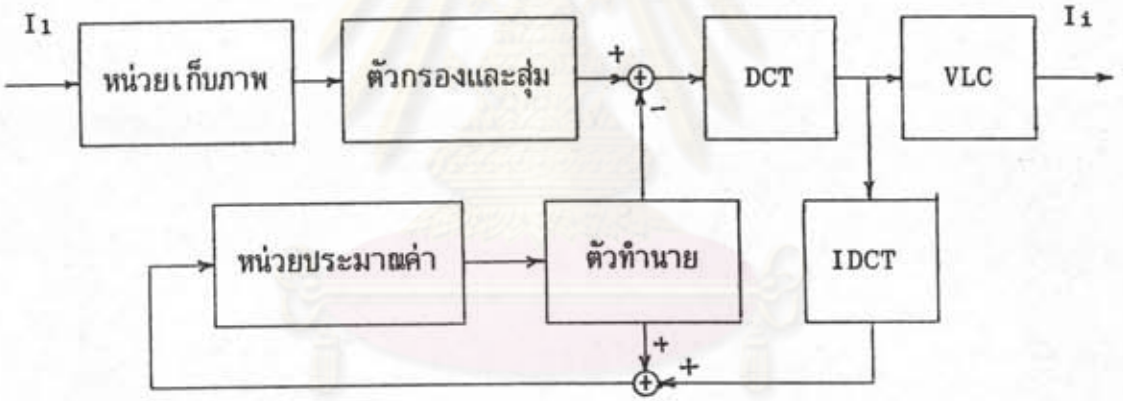


รูปที่ 2.3 แผนภูมิแสดงขั้นตอนในการรับส่งโดยใช้โครงสร้างแบบมีลำดับชั้น

ในขั้นที่สาม และ ขั้นที่สี่ ภาพต้นแบบจะถูกลบด้วยภาพที่อยู่ในตัวทำนายที่ได้จากขั้นก่อนโดยไม่มีการกรอง และ สุ่ม แล้วนำภาพแสดงความผิดพลาดที่ได้นี้ไปประมวลผลในลักษณะเดียวกับ 2 ขั้นแรก โดยให้มีอัตราข้อมูลเป็น 0.5 และ 1.5 บิตต่อจุดตามลำดับ อัตราข้อมูลรวมในขั้นที่สามจึงเท่ากับ 0.75 บิตต่อจุด และ ในขั้นที่สี่เป็น 2.25 บิตต่อจุด

ในกรณีที่ต้องการส่งแบบ Reversible scan อีกหนึ่งขั้นเป็นขั้นสุดท้าย ก็ทำได้โดยนำเอาภาพแสดงความผิดพลาด ที่ได้จากการนำภาพในตัวทำนายไปลบออกจากภาพต้นแบบไปทำ DPCM แล้วเข้ารหัสแบบ Huffman และส่งไปยังด้านรับ ซึ่งด้านรับก็จะทำการถอดรหัส และบวกภาพที่ได้กับภาพเดิมที่รับได้จากการส่งแบบโปรเกรสซีฟในตอนแรก ภาพที่ได้ในท้ายที่สุดนี้จะมีคุณภาพเหมือนกับภาพต้นแบบทุกประการ

2.4.1 <u>เครื่องส่ง</u> จะประกอบไปด้วยหน่วยย่อย ๆ ที่จะทำหน้าที่ดังแผนภูมิในรูปที่ 2.3 ซึ่งแทนได้ด้วย แผนภาพบล็อกดังรูปที่ 2.4 หน่วยย่อย ๆ เหล่านี้จะมีหน้าที่แตกต่างกัน คือ



รูปที่ 2.4 แผนภาพแสดงบล็อกที่ทำหน้าที่ต่าง ๆ ในตัวเครื่องส่ง

2.4.1.1 <u>หน่วยเก็บภาพ (Frame storage)</u> เป็นบัฟเฟอร์ที่ใช้ในการเก็บภาพที่ต้องการประมวลผลไว้ ตลอดการทำงานในขั้นต่าง ๆ

2.4.1.2 <u>ตัวกรองผ่านต่ำ และ ตัวสุ่ม (Low pass filter and Sampler)</u> ตัวสุ่มจะทำหน้าที่สุ่มภาพต้นแบบให้มีขนาดตามที่กำหนดในขั้นต่าง ๆ เพื่อลดจำนวนของอัตราข้อมูลลง ส่วนตัวกรองผ่านต่ำจำเป็นต้องมีเพื่อป้องกันการพับกลับของความถี่ (aliasing) ที่อาจเกิดขึ้นได้จากการสุ่ม ลักษณะของตัวกรองผ่านต่ำที่ใช้เป็นตัวกรองผ่านต่ำที่ถูกแนะนำโดย JPEG [10] ดังแสดงในรูปที่ 2.5

1	4	1
4	16	4
1	4	1

รูปที่ 2.5 ลักษณะของตัวกรองผ่านต่ำที่ใช้เพื่อป้องกันการพับกลับ

2.4.1.3 <u>หน่วย ADCT</u> เป็นหน่วยที่ทำหน้าที่แปลงข้อมูลที่เข้ามาด้วย DCT แล้วนำไปควอนไตซ์ แบ่งได้เป็น

1. หน่วยทรานส์ฟอร์ม ข้อมูลภาพที่เข้ามาจะถูกแบ่งออกเป็นบล็อกย่อยขนาด 8 x 8 จุดแล้วนำไปทรานส์ฟอร์มด้วย DCT 2 มิติ

2. หน่วยควอนไตซ์ เป็นส่วนที่ทำการควอนไตซ์สัมประสิทธิ์แต่ละความถี่ที่ได้จากข้อที่ 1 ด้วย step size ที่แตกต่างกันตามเมตริกซ์ค่าขีดเริ่มขององการมองเห็น (Visibility Threshold matrix) ดังรูปที่ 2.6 ซึ่งเป็นค่าที่ได้จากการทดลองทางจิตวิทยา (Psychovisual) [10] โดยที่การควอนไตซ์สำหรับแต่ละความถี่จะมีลักษณะ

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

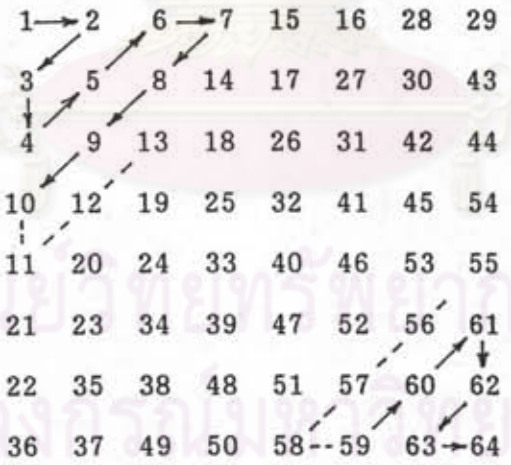
รูปที่ 2.6 เมตริกซ์แสดงค่าขีดเริ่มของการมองเห็นที่ความถี่ต่าง ๆ

แบบ Uniform ขนาดขั้นเหล่านี้สามารถเปลี่ยนแปลงค่าได้ในการควบคุมอัตราข้อมูลโดยการสเกลเมตริกซ์ข้างต้นด้วยค่าแฟกเตอร์ การควบคุมอัตราข้อมูลนี้จะกล่าวไว้ในหัวข้อที่ 2.4.3

2.4.1.4 <u>ตัวเข้ารหัสแบบแปรความยาวได้ (Variable length coder)</u> ทำหน้าที่สร้างตารางในการเข้ารหัสแบบ Huffman จากสัมประสิทธิ์ที่เหลือจากการควอนไตซ์ และเข้ารหัสสัมประสิทธิ์โดยใช้ตารางที่สร้างขึ้นมา สัมประสิทธิ์ที่เหลืออยู่จะถูกแบ่งเป็น 2 พวกในการสร้างตารางการเข้ารหัส คือ

1. สัมประสิทธิ์ DC เป็นสัมประสิทธิ์ที่แสดงถึงค่าเฉลี่ยในแต่ละบล็อกย่อย ๆ ซึ่งก็คือ ตำแหน่งที่ 1 ในรูปที่ 2.7 นั่นเอง สัมประสิทธิ์ DC นี้จะถูกนำไปทำ DPCM โดยการหาผลต่างระหว่างค่า DC ของบล็อกก่อน กับ บล็อกปัจจุบัน แล้วนำผลที่ได้ไปสร้างตารางการเข้ารหัส และ นำไปเข้ารหัสสัมประสิทธิ์ DC ต่อไป

2. สัมประสิทธิ์ AC เป็นสัมประสิทธิ์ที่แสดงถึง ขนาดของความถี่ต่าง ๆ ในแต่ละบล็อก จะถูกจัดเรียงใหม่ตามวิธีการกวาดแบบซิกแซก (zigzag scan) ตามรูปที่ 2.7 ก่อนที่จะนำไปสร้างตารางการเข้ารหัสแบบ runlength ของจำนวนศูนย์ที่ติดกันไปตามด้วยค่าสัมประสิทธิ์ที่ไม่ใช่ศูนย์ตัวแรก ตารางที่ได้จะนำไปใช้เข้ารหัสสัมประสิทธิ์ AC โดยเฉพาะ



รูปที่ 2.7 ตำแหน่งการเรียงข้อมูลตามวิธี zigzag scan

2.4.1.5 <u>หน่วย IDCT</u> จะทำการดีควอนไตซ์ (Dequantize) และทำการทรานส์ฟอร์มกลับค่าสัมประสิทธิ์ที่ได้จากหน่วย ADCT ก่อนที่จะถูกนำไปเข้ารหัสส่งไปยังด้านรับ

2.4.1.6 <u>ตัวประมาณค่า (Interpolater)</u> เป็นตัวประมาณค่าแบบเชิงเส้น (bilinear) ซึ่งจะทำให้ภาพที่ได้มีขนาดใหญ่ขึ้นสองเท่าของภาพเดิม ตัวประมาณค่าที่ใช้จะมีลักษณะดังรูปที่ 2.8

a	x2	b
x5	x1	x3
d	x4	c

รูปที่ 2.8 ลักษณะของการประมาณค่า โดยที่

a,b,c,d : จุดในรูปที่สุ่ม

x1 - x5 : จุดที่ต้องการ interpolate

ค่า x1 จนถึง x5 ในรูปที่ 2.8 สามารถประมาณได้จาก [10]

$$x1 = (a + b + c + d) / 4$$

$$x2 = (a + b) / 2$$

$$x3 = (b + c) / 2$$

$$x4 = (c + d) / 2$$

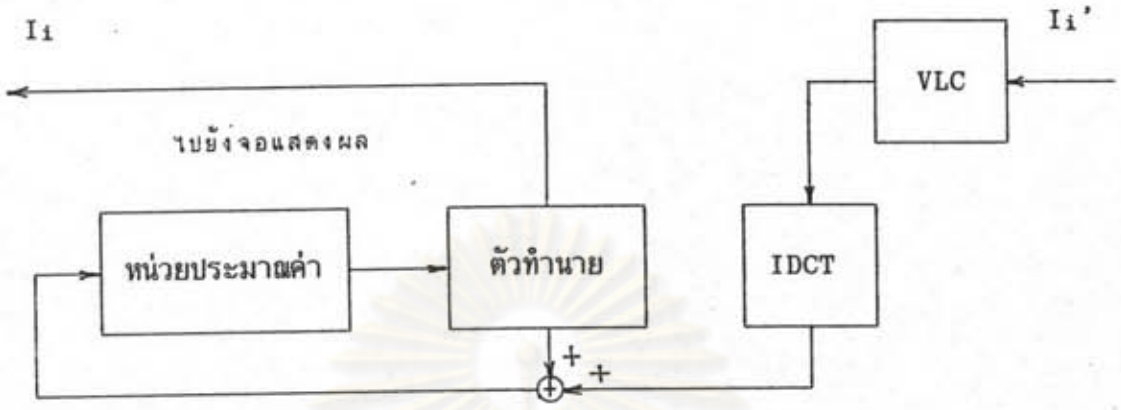
$$x5 = (d + a) / 2$$

2.4.1.7 <u>ตัวทำนาย (Predictor)</u> ทำหน้าที่เป็นบัฟเฟอร์ที่เก็บภาพซึ่งด้านรับควรรับได้ ค่าในตัวทำนายนี้จะเอาไปลบออกจากภาพต้นแบบที่ขั้นต่าง ๆ ซึ่งทำให้แก้ไขความผิดเพี้ยนที่เกิดขึ้นได้ในขั้นถัดมา

2.4.1.8 <u>หน่วยย่อยอื่น ๆ</u> ได้แก่ หน่วยที่ทำหน้าที่ทางคณิตศาสตร์อื่น ๆ เช่น ตัวบวก และ ตัวลบ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

2.4.2 <u>เครื่องรับ</u> เครื่องรับก็จะประกอบด้วยหน่วยย่อย ๆ ที่ทำหน้าที่ คล้ายกับเครื่องส่ง แต่จะมีความซับซ้อนน้อยกว่า ดังแสดงในรูปที่ 2.9 คือ



รูปที่ 2.9 แผนภาพแสดงบล็อกที่ทำหน้าที่ต่าง ๆ ในตัวเครื่องรับ

2.4.2.1 <u>ตัวถอดรหัสแบบแปรความยาวได้ (Variable length decoder)</u> จะทำหน้าที่ถอดรหัสข้อมูลที่ได้มาจากช่องสัญญาณ โดยใช้ตารางการเข้ารหัสซึ่งจะถูกส่งมาด้วยในตอนต้น ๆ ของการส่งข้อมูล

2.4.2.2 <u>หน่วย IDCT</u> จะทำหน้าที่เหมือนกับทางด้านส่ง

2.4.2.3 <u>ตัวประมาณค่า</u> จะมีลักษณะเป็นเชิงเส้นเหมือนด้านส่ง แต่จะมีข้อแตกต่างกันเล็กน้อย คือ จะทำการประมาณค่าให้ได้ขนาดของภาพเท่ากับขนาดของภาพต้นแบบ เพื่อนำไปแสดงผลทางจอภาพได้โดยตรง

2.4.2.4 <u>หน่วยเก็บภาพ</u> เป็นบัฟเฟอร์ไว้เก็บภาพที่จะส่งไปแสดงผล และ บวกเข้ากับข้อมูลที่จะได้มาในขั้นถัดไปเพื่อปรับปรุงภาพให้ดีขึ้น

2.4.3 <u>การควบคุมอัตราข้อมูล</u> อัตราข้อมูลของภาพในขั้นต่าง ๆ จะควบคุมผ่านทางหน่วยควอนไตซ์ และ ตัวเข้ารหัสแบบแปรความยาวได้ โดยการสเกลเมตริกซ์ค่าขีดเริ่มของการมองเห็นด้วยค่าที่เรียกว่า ค่าแฟกเตอร์ขีดเริ่ม (threshold factor) ซึ่งจะขึ้นอยู่กับความซับซ้อนของภาพที่ถูกประมวลผล ปกติภาพสองภาพที่มีคุณภาพทาง Psychovisual พอ ๆ กันไม่จำเป็นต้องมีอัตราข้อมูลเท่ากัน การควบคุมอัตราข้อมูลให้เท่ากันจึงต้องใช้ค่าแฟกเตอร์ขีดเริ่มซึ่งแตกต่างกันไป การหาค่าแฟกเตอร์ขีดเริ่มของแต่ละภาพจะหาได้จาก ความสัมพันธ์ระหว่างอัตราข้อมูล B ที่ได้จากการเข้ารหัส กับ ค่าแฟคเตอร์ขีดเริ่ม F ที่ได้มาโดยวิธี

017717

เอมไพริกัล (Empirical Method) [10] คือ

$$\log(B) = a * \log(F) + b$$

โดย a และ b เป็นค่าคงที่ขึ้นกับลักษณะของภาพนั้น ๆ และ F เป็นค่าแฟกเตอร์ขีดเริ่มที่ถูกสเกลขึ้นด้วย 50 การควบคุมอัตราข้อมูลให้ได้ตามที่ต้องการจะใช้วิธีวนซ้ำ (iterate) โดยใช้วิธีของ Newton-Ralphson จนกว่าจะได้ค่าแฟกเตอร์ขีดเริ่มที่จะทำให้ได้อัตราข้อมูลจากการเข้ารหัสเท่ากับอัตราข้อมูลตามที่ต้องการ หรือ ใกล้เคียงที่สุด โดยปกติค่าเริ่มต้น (initial value) ที่พอเหมาะ คือ a = -2/3 และ F = 50 [10]

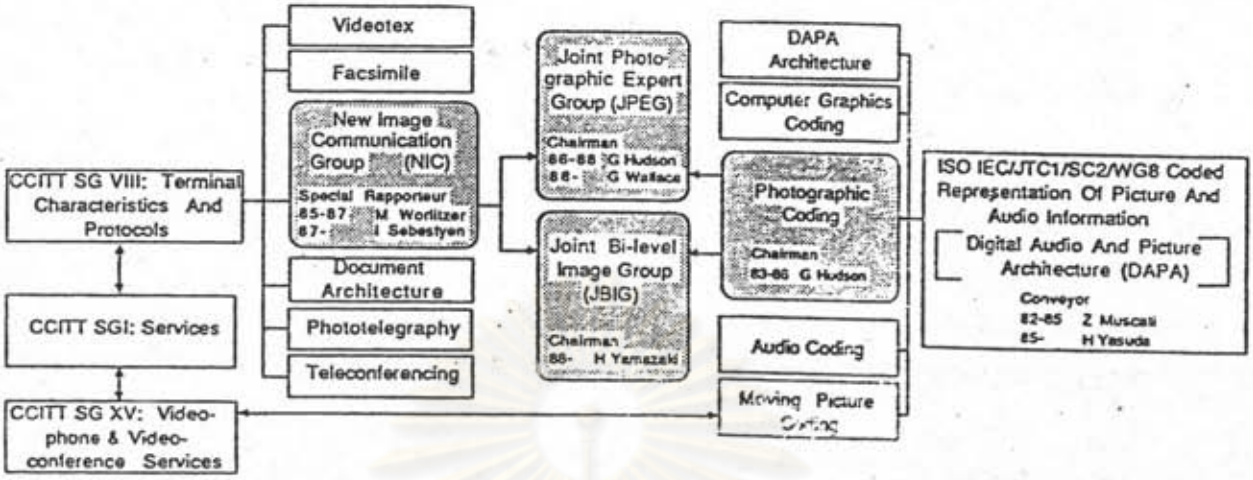
2.4.4 <u>รูปแบบ (format) ของการส่งข้อมูล</u> ตามที่ได้กล่าวมาแล้วว่าการส่งแบบโปรเกรสซีฟจำเป็นต้องส่งบิต overhead กำกับไปด้วย เพื่อทำให้การประมวลผลภาพกลับเป็นไปอย่างถูกต้อง บิต overhead เหล่านี้จะประกอบด้วยค่าแฟกเตอร์ขีดเริ่ม (threshold factor), จำนวนบล็อกในขั้นต่าง ๆ , ตารางการเข้ารหัสสำหรับสัมประสิทธิ์ DC และ AC ค่าต่าง ๆ เหล่านี้จะถูกอัด (pack) เข้าด้วยกันในตอนต้นของการส่ง โดยรูปแบบที่ใช้ในการส่งแสดงได้ดังรูปที่ 2.10

แฟกเตอร์ค่าขีดเริ่ม	จำนวนบล็อก	ตารางการเข้ารหัสของ DC	DC	ตารางการเข้ารหัสของ AC	AC

รูปที่ 2.10 รูปแบบที่ใช้ในการรับส่งข้อมูลในแต่ละขั้น

2.5 <u>มาตรฐานของ JPEG กับ การส่งแบบโปรเกรสซีฟ</u>

JPEG (Joint Photographic Expert Group) เป็นกลุ่มผู้เชี่ยวชาญทางด้านการประมวลผลภาพนิ่งซึ่งจัดตั้งขึ้นมาจากการร่วมมือระหว่าง CCITT และ ISO พร้อมกับกลุ่มผู้เชี่ยวชาญอื่นๆ ในปี 1986 โดยมีหน้าที่กำหนดมาตรฐาน และข้อแนะนำสำหรับการประมวลผลภาพนิ่งโดยเฉพาะ [11]



รูปที่ 2.11 กลุ่มผู้เชี่ยวชาญที่เกิดจากความร่วมมือระหว่าง CCITT กับ ISO

JPEG ได้ทดลองวิธีต่าง ๆ เพื่อใช้เป็นวิธีมาตรฐาน ในปลายปี 1990 ก็ได้ออกมาตรฐานฉบับร่างออกมา [12] โดยกำหนดระบบพื้นฐาน (Based-line System) ไว้สำหรับการประมวลผลและส่งภาพนิ่ง ซึ่งมาตรฐานฉบับร่างมีหัวข้อสำคัญดังนี้

1. เทคนิคการลดข้อมูล (Data compressing techniques) มีการกล่าวถึงเทคนิค 2 อย่าง คือ วิธีการทรานส์ฟอร์ม DCT ซึ่งใช้ในระบบพื้นฐาน และ วิธี DPCM ซึ่งมุ่งหมายให้ใช้ในขั้น lossless transmission

2. การเข้ารหัสประเภทเอนโทรปี นอกจากจะใช้การเข้ารหัสแบบ Huffman เป็นวิธีหลักแล้ว JPEG ยังให้เลือกใช้วิธีการเข้ารหัสแบบ Adaptive Binary Arithmetic Coding (ABAC) ซึ่งเป็น การเข้ารหัสแบบเอนโทรปีที่มีประสิทธิภาพมากอีกแบบหนึ่งได้อีกด้วย

3. โหมดปฏิบัติงาน (mode of operation) JPEG ได้กำหนดโหมดการปฏิบัติงานไว้ 2 โหมด คือ การส่งในโหมดซีเควนเชียล (Sequential) และ การส่งในโหมดโปรเกรสซีฟ

4. การส่งในโหมดโปรเกรสซีฟ มีวิธีพื้นฐาน 2 วิธี คือ วิธีประมาณค่าแบบต่อเนื่อง (successive approximation) และ วิธีเลือกความถี่ที่จะส่ง (spectral selection)

5. โครงสร้างแบบมีลำดับชั้น ใช้สำหรับการส่งในโหมดโปรเกรสซีฟ ในกรณีที่มีแบนด์วิดท์ในการส่งจำกัด ซึ่งวิทยานิพนธ์ฉบับนี้ก็ได้ใช้โครงสร้างนี้ในการสร้างระบบ