



บทที่ 11 : การบีบอัดข้อมูล (Data Compression)

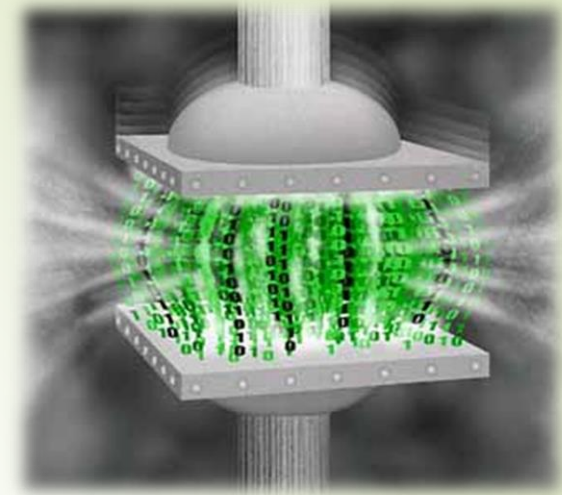
สร212 ระบบสื่อประสมสำหรับธุรกิจ

อาจารย์อภิพงศ์ ปิงยศ

apipong.ping@gmail.com

Outline

- ทำความรู้จักกับการบีบอัดข้อมูล
- ประเภทของการบีบอัดข้อมูล
- วิธีการบีบอัดข้อมูลภาพนิ่งและวิดีโอ
- เทคนิคการบีบอัดข้อมูล
- มาตรฐานการบีบอัดข้อมูลแบบ JPEG
- มาตรฐานการบีบอัดข้อมูลแบบ MPEG
- ซอฟต์แวร์สำหรับบีบอัดวิดีโอ

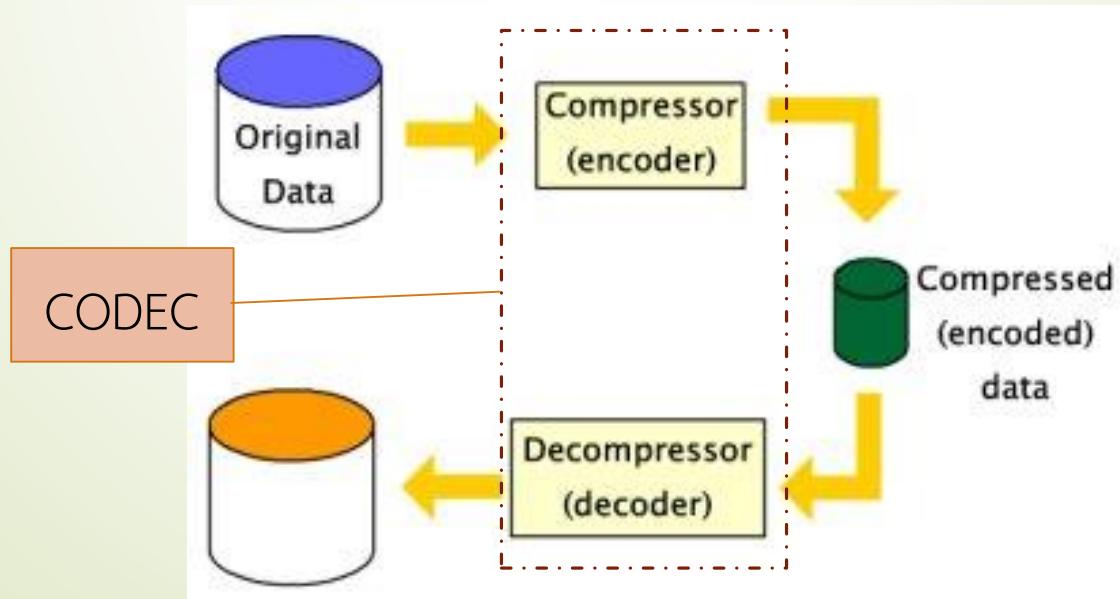


ทำความรู้จักกับการบีบอัดข้อมูล

- ข้อมูลดิบ (Raw Data) ที่ได้จากเครื่องมือต่างๆนั้นมีขนาดของไฟล์ที่ใหญ่ ซึ่งใช้พื้นที่ในการจัดเก็บมาก จึงจำเป็นต้องมีการบีบอัด (Compression) เพื่อลดขนาดของข้อมูลลง
- **การบีบอัดข้อมูล** คือกระบวนการเข้ารหัสข้อมูลที่เลือกเอาเฉพาะบิตที่จำเป็นเท่านั้น เพื่อประหยัดพื้นที่ในการจัดเก็บ และสามารถรับส่งข้อมูลผ่านเครือข่ายได้รวดเร็วยิ่งขึ้น

ทำความรู้จักกับการบีบอัดข้อมูล [2]

- CODEC (Coder-Decoder) เป็นกลไกสำหรับเข้ารหัสและถอดรหัสข้อมูลดิจิทัล เมื่อเข้ารหัสด้วย CODEC ชนิดใด ก็จะต้องถอดรหัสด้วย CODEC ชนิดเดียวกัน



ทำความรู้จักกับการบีบอัดข้อมูล [3]

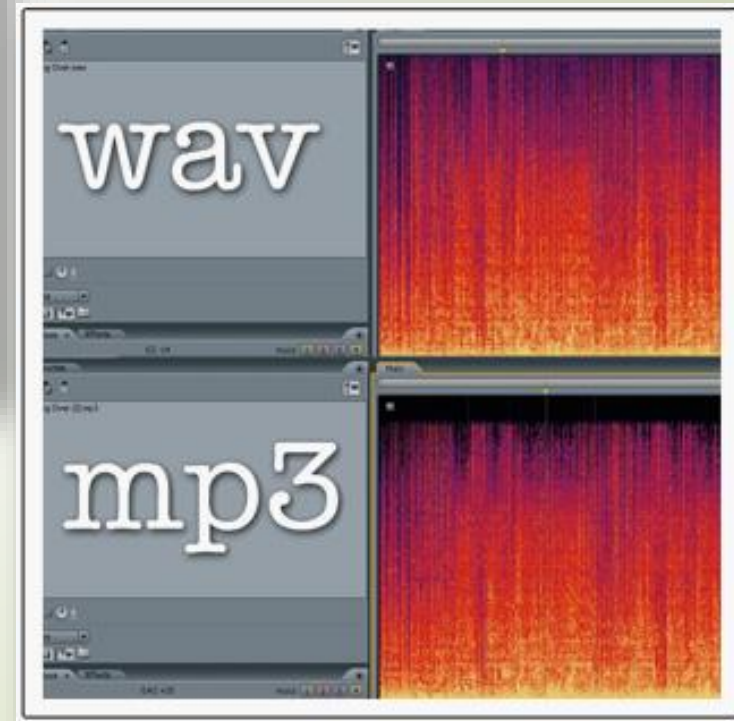
- CODEC จะใช้อัลกอริทึมในการหาข้อมูลที่มีความซ้ำซ้อนกัน (Redundant) และลดความซ้ำซ้อนนั้นลง โดยไม่ส่งผลกระทบต่อคุณภาพของข้อมูล ข้อมูลที่ถูกบีบอัดแล้ว ก็จะมีนามสกุลของไฟล์แตกต่างกันไปตามวิธีการบีบอัด



ประเภทของการบีบอัดข้อมูล

- **Lossless Compression** ใช้อัลกอริทึมในการบีบอัดข้อมูลที่ไม่ทำให้ข้อมูลสูญหายในการบีบอัด ข้อมูลจึงมีความสมบูรณ์เหมือนต้นฉบับ แต่ก็ใช้พื้นที่ในการจัดเก็บมาก ตัวอย่างเช่น PNG, TIFF, GIF, AVI
- **Lossy Compression** จะตัดข้อมูลบางส่วนออกไปเพื่อลดขนาดของไฟล์ โดยข้อมูลที่ซ้ำซ้อนจะถูกตัดทิ้งอย่างถาวร ข้อดีคือขนาดข้อมูลจะลดลงมาก แต่ข้อเสียคือข้อมูลจะไม่ละเอียดเหมือนต้นฉบับ

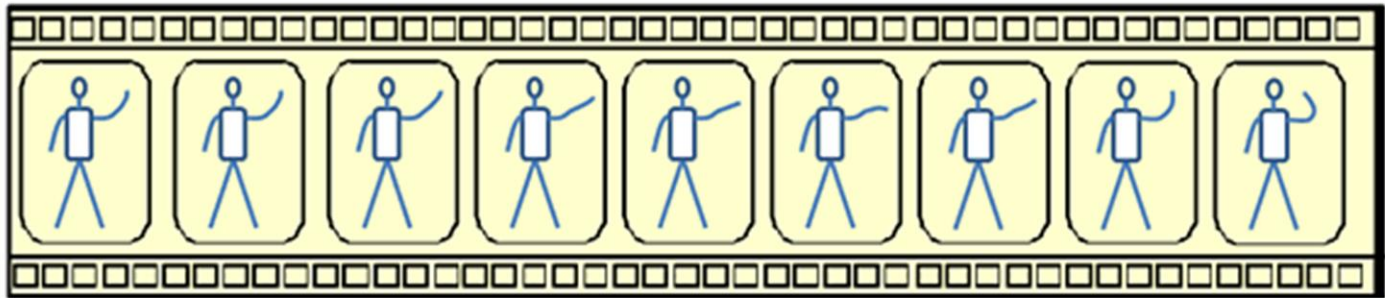
Lossless vs Lossy Compression



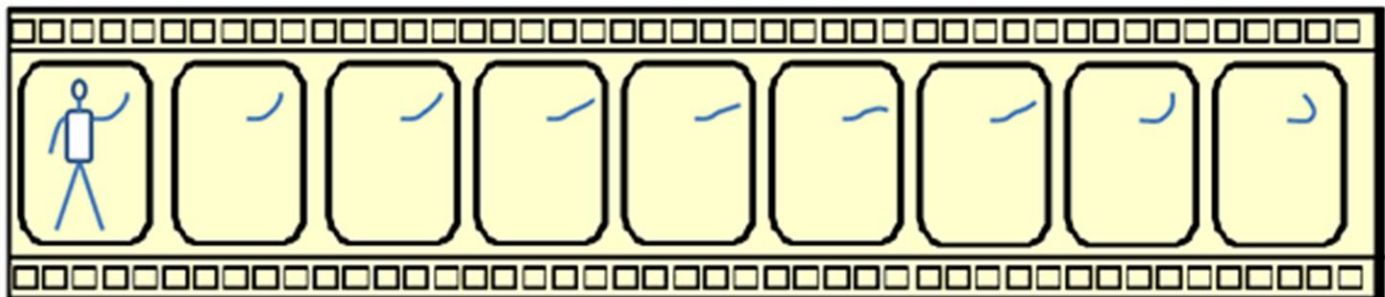
วิธีการบีบอัดข้อมูลภาพนิ่งและวิดีโอ

- ▶ **Intraframe Compression** นิยมใช้กับภาพนิ่ง โดยการลดความซ้ำซ้อนเชิงพื้นที่ (Spatial Redundancy) ที่มีคุณสมบัติใกล้เคียงกัน เช่น พิกเซลในรูปที่มีค่าสีเหมือนกัน เป็นต้น
- ▶ **Interframe Compression** นิยมใช้กับไฟล์วิดีโอ ใช้วิธีการคำนวณความซ้ำซ้อนของข้อมูลระหว่างเฟรมที่อยู่ติดกัน ที่เรียกว่าความซ้ำซ้อนเชิงเวลา (Temporal Redundancy)

Intraframe vs Interframe Compression

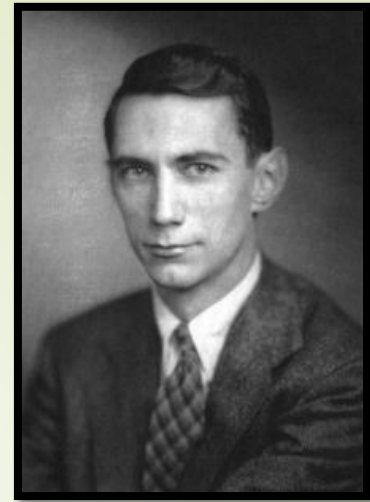


Intraframe compression
Every frame is encoded individually



Interframe compression
Only the differences between frames are encoded
for each group of frames

เทคนิคการบีบอัดข้อมูล



Lossless Compression

- ▶ Entropy ใช้แบบจำลองความน่าจะเป็นในทางสถิติคาดเดาว่าเหตุการณ์ถัดไปคืออะไร โดยใช้อัลกอริทึม Shannon's Law of Entropy
- ▶ RLE (Run Length Encoding)
- ▶ Huffman Coding
- ▶ Arithmetic Coding
- ▶ LZW (Lempel-Ziv-Welsh Coding)
- ▶ Etc...

เทคนิคการบีบอัดข้อมูล [2]

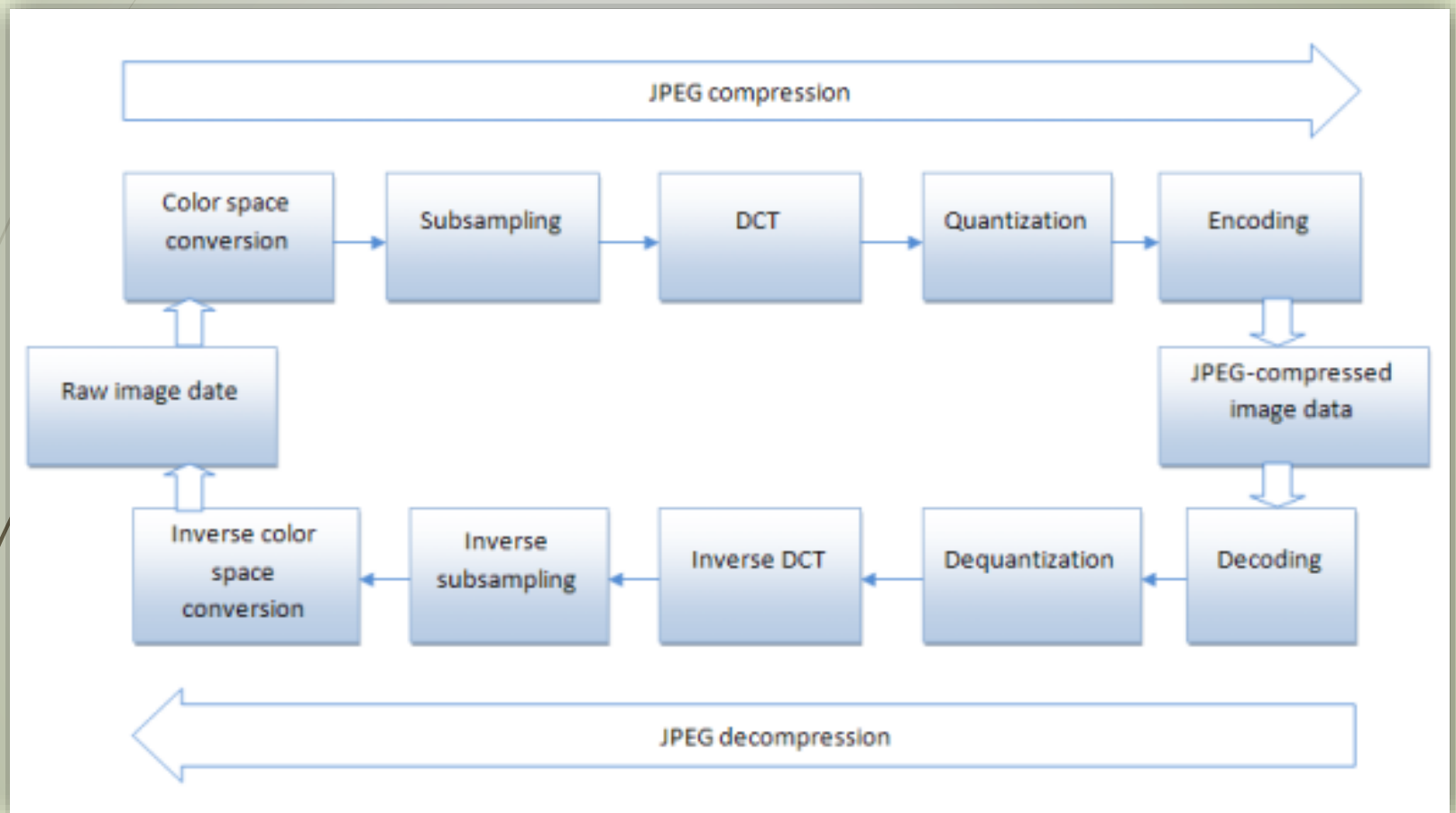
Lossy

- Transform Coding
- Psycho-Analysis เข้ามหัสเสียงตามการได้ยินของมนุษย์
- Interframe Correlation จะใช้อัลกอริทึม JPEG Compression มาทำการบีบอัดวิดีโอ
- Fractal

มาตรฐานการบีบอัดข้อมูลแบบ JPEG

- ▶ JPEG (Joint Photographers Expert Group) เป็นมาตรฐานการบีบอัดภาพที่ใช้เทคนิคหลายรูปแบบมาผสมผสานกัน เช่น RLE และ Huffman Coding เป็นต้น
- ▶ รองรับการบีบอัดทั้งแบบ Lossless และ Lossy
- ▶ ภาพที่ได้จากการบีบอัดจะเป็นภาพแบบ Lossy

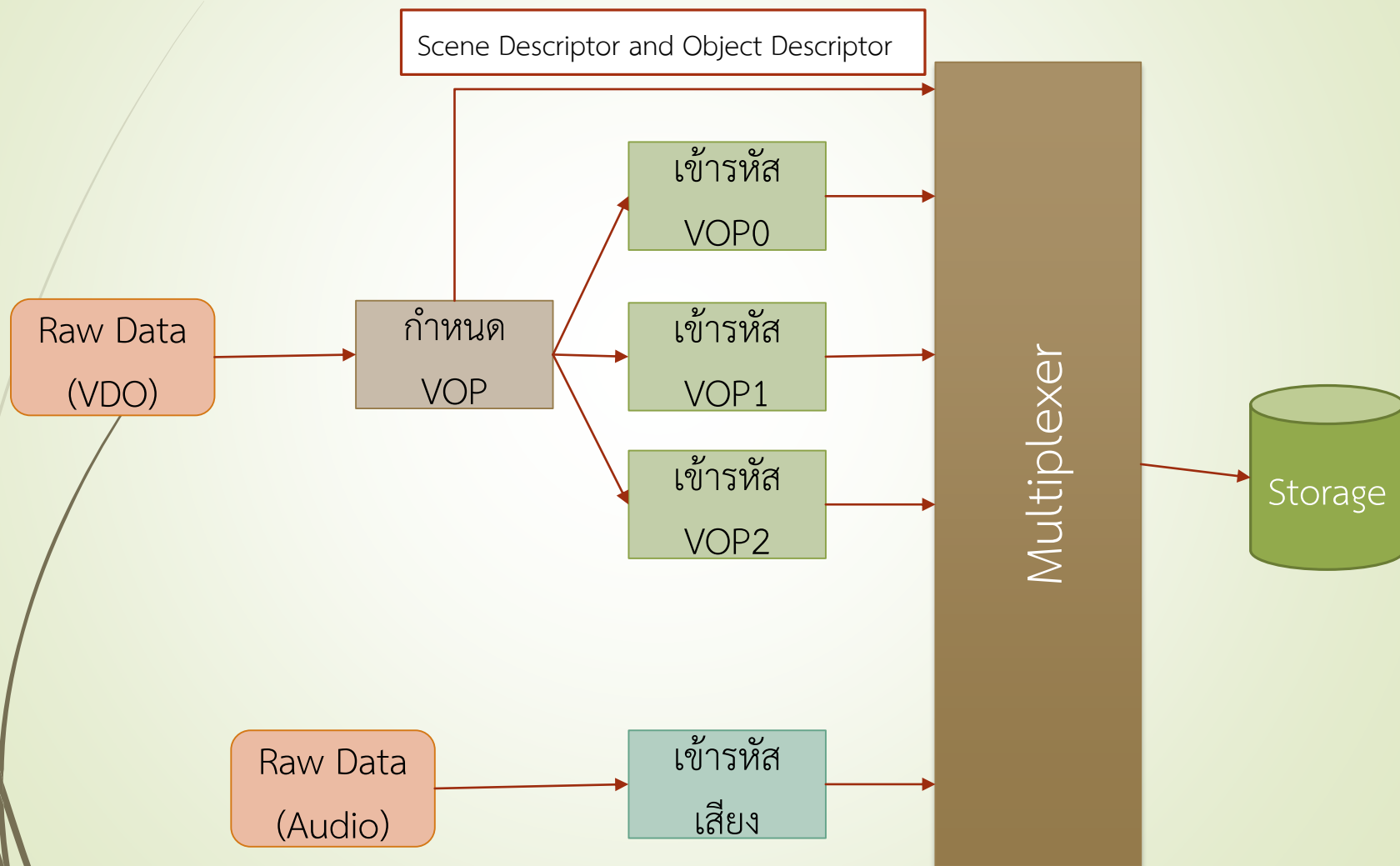
ผังการทำงานของ การบีบอัดข้อมูลแบบ JPEG



มาตรฐานการบีบอัดข้อมูลแบบ MPEG

- ▶ MPEG (Motion Picture Expert Group) เป็นมาตรฐานการบีบอัดข้อมูลเสียงและวิดีโอ วิธีการคือจะหาข้อมูลที่ซ้ำซ้อนเพื่อนำมาคำนวณและแทนค่าจุดสีต่างๆ ประกอบไปด้วย
 - ▶ MPEG-1 Audio (MP1, MP2, MP3) และ MPEG-1 Video
 - ▶ MPEG-2 Audio และ MPEG-2 Video
 - ▶ MPEG-4 จะมีกระบวนการนำเสียงหรือวิดีโอมากกว่าหนึ่งออบเจกต์มารวมกันเป็นออบเจกต์ใหม่ (คล้ายการซ้อน Layer ในโปรแกรมตัดต่อรูปภาพ) เรียกแต่ละออบเจกต์ว่า Video Object Plane (VOP)

MPEG-4 Encoder



ซอฟต์แวร์สำหรับบีบอัดวิดีโอ

