Lossy Image Compression Methods

- DCT Compression
  - JPEG
- Wavelet Compression
  - SPIHT
  - UWIC (University of Washington Image Coder)
  - EBCOT (JPEG 2000)
- Scalar quantization (SQ).
- Vector quantization (VQ).

JPEG Standard

- JPEG - Joint Photographic Experts Group
- JPEG 2000 uses wavelet compression.

Barbara

32:1 compression ratio
0.25 bits/pixel (8 bits)

VQ

JPEG

VQ
Images and the Eye

- Images are meant to be viewed by the human eye (usually).
- The eye is very good at "interpolation", that is, the eye can tolerate some distortion. So lossy compression is not necessarily bad. The eye has more acuity for luminance (gray scale) than chrominance (color).
  - Gray scale is more important than color.
  - Compression is usually done in the YUV color coordinates. Y for luminance and U, V for color.
  - U and V should be compressed more than Y
  - This is why we will concentrate on compressing gray scale (8 bits per pixel) images.

Distortion

\[ MSE = \frac{1}{N} \sum_{i=1}^{N} (x_i - \hat{x}_i)^2 \]

PSNR

- Peak Signal to Noise Ratio (PSNR) is the standard way to measure fidelity.

\[ PSNR = 10 \log_{10} \left( \frac{m^2}{MSE} \right) \]

where \( m \) is the maximum value of a pixel possible.
For gray scale images (8 bits per pixel) \( m = 255 \).

- PSNR is measured in decibels (dB).
  - .5 to 1 dB is said to be a perceptible difference.
  - Decent images start at about 30 dB

Rate-Fidelity Curve

- Properties:
  - Increasing
  - Slope decreasing
PSNR is not Everything

\[
\text{PSNR} = 25.8 \text{ dB}
\]

PSNR Reflects Fidelity (1)

\[
\begin{align*}
\text{PSNR} & = 25.8 \text{ dB} \\
\text{12.8 : 1}
\end{align*}
\]

PSNR Reflects Fidelity (2)

\[
\begin{align*}
\text{PSNR} & = 24.2 \text{ dB} \\
\text{25.6 : 1}
\end{align*}
\]

PSNR Reflects Fidelity (3)

\[
\begin{align*}
\text{PSNR} & = 23.2 \text{ dB} \\
\text{51.2 : 1}
\end{align*}
\]

Idea of Transform Coding

- Transform the input pixels \(x_0, x_1, \ldots, x_{N-1}\) into coefficients \(c_0, c_1, \ldots, c_{N-1}\) (real values)
  - The coefficients are have the property that most of them are near zero
  - Most of the "energy" is compacted into a few coefficients
- Quantize the coefficients
  - This is where there is loss, since coefficients are only approximated
  - Important coefficients are kept at higher precision
- Entropy encode the quantization symbols

Decoding

- Entropy decode the quantized symbols
- Compute approximate coefficients \(c'_0, c'_1, \ldots, c'_{N-1}\) from the quantized symbols.
- Inverse transform \(c'_0, c'_1, \ldots, c'_{N-1}\) to \(x'_0, x'_1, \ldots, x'_{N-1}\) which is a good approximation of the original \(x_0, x_1, \ldots, x_{N-1}\).
Block Diagram of Transform Coding

Mathematical Properties of Transforms

- Linear Transformation - Defined by a real n x n matrix \( A = (a_{ij}) \)

\[
\begin{bmatrix}
a_{00} & \cdots & a_{0N-1} \\
\vdots & \ddots & \vdots \\
a_{N-1,0} & \cdots & a_{N-1,N-1}
\end{bmatrix}
\begin{bmatrix}
x_0 \\
\vdots \\
x_{N-1}
\end{bmatrix}
= \begin{bmatrix}
c_0 \\
\vdots \\
c_{N-1}
\end{bmatrix}
\]

- Orthonormality \( A^{-1} = A^T \) (transpose)

Why Coefficients

Why Orthonormality

- The energy of the data equals the energy of the coefficients

\[
\sum_{i=0}^{N-1} c_i^2 = c^T c = (Ax)^T (Ax)
= (x^T A^T) (Ax) = x^T (A^T A) x = x^T x = \sum_{i=0}^{N-1} x_i^2
\]

Squared Error is Preserved with Orthonormal Transformations

- In lossy coding we only send an approximation \( c'_i \) of \( c_i \) because it takes fewer bits to transmit the approximation.

Let \( c_i = c'_i + \epsilon_i \)

\[
\sum_{i=0}^{N-1} \epsilon_i^2 = \sum_{i=0}^{N-1} (c_i - c'_i)^2 = (c - c')^T (c - c') = (Ax - Ax')^T (Ax - Ax')
= (A(x - x'))^T A(x - x') = x^T (A^T A x - x')^T (x - x')
= x^T (x - x')^T (x - x')
= \sum_{i=0}^{N-1} (x_i - x'_i)^2
\]

Squared error in original.

Compaction Example

\[
A = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}
\]

\( A^2 = A \Rightarrow A^{-1} = A \)

\( A^T = A^{-1} \) orthonormal

\[
\left[ \begin{array}{c}
1 \\
1 \\
\end{array} \right] = \left[ \begin{array}{c}
\sqrt{3} b \\
0 \\
\end{array} \right]
\] compaction
Discrete Cosine Transform

\[ d_i = \begin{cases} \frac{1}{\sqrt{N}} & i = 0 \\ \frac{2}{\sqrt{N}} \cos \left( \frac{(2i+1)\pi}{2N} \right) & i > 0 \end{cases} \]

\( N = 4 \)

\[
D = \begin{bmatrix}
0.5 & 0.5 & 0.5 & 0.5 \\
0.270598 & -0.270598 & -0.653281 & 0.653281 \\
0.5 & 0.5 & 0.5 & 0.5 \\
0.270598 & -0.653281 & 0.653281 & -0.270598 \\
\end{bmatrix}
\]

Basis Vectors

Decomposition in Terms of Basis Vectors

Block Transform

2-Dimensional Block Transform

Block of pixels \( X \)

Transform \( A \)

\[
A = \begin{bmatrix}
a_{00} & a_{01} & a_{02} & a_{03} \\
\vdots & \vdots & \vdots & \vdots \\
a_{30} & a_{31} & a_{32} & a_{33} \\
\end{bmatrix}
\]

Transform rows \( r_i = \sum_{j=0}^{3} a_{ij} x_k \)

Transform columns \( c_j = \sum_{i=0}^{3} a_{ij} x_k \)

Summary \( C = AXA^T \)

8x8 DCT Basis
Importance of Coefficients

- The DC coefficient is the most important.
- The AC coefficients become less important as they are farther from the DC coefficient.
- Example Bit Allocation

```
0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 1
0 0 0 0 0 0 1 2
0 0 0 0 0 1 2 3
0 0 0 0 1 2 3 5
0 0 0 1 2 3 5 7
0 0 1 2 3 5 7 8
```

Compression: 55 bits for 64 pixels = .86 bpp

Quantization

- For a nxn block we construct a nxn matrix $Q$ such that $Q_{ij}$ indicates how many quantization levels to use for coefficient $c_{ij}$.
- Encode $c_{ij}$ with the label:
  
  $$s_{ij} = \frac{c_{ij} + 0.5}{Q_{ij}}$$
  
  Larger $Q_{ij}$ indicates fewer levels.
- Decode $s_{ij}$ to
  
  $$c'_{ij} = s_{ij} Q_{ij}$$

Example Quantization

- $c = 54.2$, $Q = 24$
  
  $$s = \frac{54.2}{24} + 0.5 = 2$$
  
  $$c' = 2\cdot24 + 48$$
- $c = 54.2$, $Q = 12$
  
  $$s = \frac{54.2}{12} + 0.5 = 5$$
  
  $$c' = 5\cdot12 + 60$$
- $c = 54.2$, $Q = 6$
  
  $$s = \frac{54.2}{6} + 0.5 = 9$$
  
  $$c' = 9\cdot6 + 54$$

Example Quantization Table

```
16 11 10 16 24 40 55 61
12 14 24 46 12 30 60 55
14 16 24 48 57 59 56
14 17 22 29 57 67 79 62
18 35 37 35 88 109 133 149
24 35 45 54 69 78 104 113 92
49 64 87 121 120 101
72 92 98 112 100 103 99
```

Increase the bit rate = halve the table
Decrease the bit rate = double the table

Zig-Zag Coding

- DC label is coded separately.
- AC labels are usually coded in zig-zag order using a special entropy coding to take advantage of the ordering of the bit allocation (quantization).

```
1 2 3 4 5 6 7 8
9 10 11 12 13 14 15 16
17 18 19 20 21 22 23 24
25 26 27 28 29 30 31 32
33 34 35 36 37 38 39 40
41 42 43 44 45 46 47 48
49 50 51 52 53 54 55 56
57 58 59 60 61 62 63 64
```

JPEG (1987)

- Let $P = [p_{ij}]$, $0 < i, j < N$ be an image with $0 < p_{ij} < 256$.
- Center the pixels around zero
  
  $$x_{ij} = p_{ij} - 128$$
- Code 8x8 blocks of $P$ using DCT
- Choose a quantization table.
  
  - The table depends on the desired quality and is built into JPEG
- Quantize the coefficients according to the quantization table.
  
  - The quantization symbols can be positive or negative.
- Transmit the labels (in a coded way) for each block.
Block Transmission

- DC coefficient
  - DC coefficients don't change much from block to neighboring block. Hence, their labels change even less.
  - Predictive coding using differences is used to code the DC label.
- AC coefficients
  - Do a zig-zag coding.

Example Block of Labels

<table>
<thead>
<tr>
<th>Z</th>
<th>C</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>00</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1100</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>11101</td>
<td>111</td>
</tr>
</tbody>
</table>

Coding order of AC labels

\[ 2 \rightarrow \quad 3 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad \ldots \]

Coding Labels

- Categories of labels
  - 1 \{0\}
  - 2 \{-1, 1\}
  - 3 \{-3, -2, 2, 3\}
  - 4 \{-7, -6, -5, -4, 4, 5, 6, 7\}
- Label is indicated by two numbers C,B
- Examples
  - label C,B
    - 0 1
    - 2 3, 2
    - -4 4, 3

Coding AC Label Sequence

- A symbol has three parts (Z,C,B)
  - Z for number of zeros preceding a label \(0 \leq Z \leq 15\)
  - C for the category of the label
  - B for a C-1 bit number for the actual label
- End of Block symbol (EOB) means the rest of the block is zeros. EOB = (0, 0, -)
- Example:
  \[ 2 \rightarrow -8 \quad 3 \quad 0 \quad 0 \quad 0 \quad 1 \quad 1 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad \ldots \]

Notes on Transform Coding

- Video Coding
  - MPEG – uses DCT
  - H.263, H.264 – uses DCT
- Audio Coding
  - MP3 = MPEG 1- Layer 3 uses DCT
- Alternative Transforms
  - Lapped transforms remove some of the blocking artifacts.
  - Wavelet transforms do not need to use blocks at all.