JPEG should have

- parameterized compression / quality tradeoff
- general applications
- tractable enc/dec
- lossless version available

Codec { encoder, decoder

Some psychophysical background...

- Just noticeable difference (JND)

  For any sensory system S, let I be the magnitude of stimulus intensity and ΔI be the difference threshold s.t. $Pr(S(I \neq I + \Delta I)) \geq 1/2$.

  We have $\frac{\Delta I}{I} = k_s$, a constant depending on S.

  ▷ $\frac{\Delta I}{I}$ = "Weber's fraction"

  ▷ $I + \Delta I$ = "JND" w.r.t. I.

  cf) Weber's fraction fails for extreme I.

- Fechner's law

  The magnitude of the sensation is a logarithmic function of the stimulus : $M = k_s \log I$.

* H.R. Schiffman. "Sensation and Perception: An Integrated approach."

- Luminance vs. Chrominance

  휘도   색차

  ↳ Recognized by "rods", sensitive (100M) (Y)

  ↳ Recognized by "cones", not as sensitive (6M)

  (Cb: Chrominance Blue; $B' - Y'$) (Cr: Chrominance red)

Alg JPEG.

Given an $N \times M$ image:

Convert image to $YCbCr$ color space

Split image into 8×8 blocks

FOR each block B do:

  Perform $\boxed{DCT}$ on B to obtain 64 coeffs. ← Some transformation

  $\boxed{\text{"Quantization"}}$ ← compression

  Encode result using run-length enc & Huffman.

end FOR.

return encoded string.

Karhunen-Loève transform

Consider a vector $\vec{x}$ drawn from some pdf F.

Let $\vec{m}_x := E[\vec{x}]$.

The <u>Covariance matrix</u> of the population is

$$
\begin{aligned}
C_{\vec{x}} &= E[(\vec{x}-\vec{m}_x)(\vec{x}-\vec{m}_x)^T] \\
&= E[\vec{x}\vec{x}^T] - \vec{m}_x\vec{m}_x^T \\
&\underset{\text{large } M}{\approx} \frac{1}{M}\sum_{k=1}^{M} \vec{x}_k \vec{x}_k^T - \vec{m}_x \vec{m}_x^T \quad \cdots \text{ law of large numbers}
\end{aligned}
$$

Let $(\lambda_1, \vec{e}_1), (\lambda_2, \vec{e}_2), \dots, (\lambda_n, \vec{e}_n)$ be
n ↳ "dimension" of $\vec{x}$

(eigen value, eigen vector) pairs of $C_{\vec{x}}$.

Define transform T:

$$
A = \begin{bmatrix} - \vec{e}_1 - \\ - \vec{e}_2 - \\ \vdots \\ - \vec{e}_n - \end{bmatrix}
$$

$$
T(\vec{x}) = A(\vec{x} - \vec{m}_x) \quad \text{"Karhunen-Loève"}
$$

- $E[T(\vec{x})] = 0$

- $C_{\vec{y}}$ is a diagonal matrix w/ eigenvals for C_x on diagonal
  ↳ $\vec{y} = T(\vec{x})$ $\Rightarrow$ coefficients for $\vec{y}$ are independant.
- W.L.O.G. assume $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ and use only top k eigenvectors for A. Recon

Note that $A^T = A^{-1}$ (orthogonal)

$\rightsquigarrow$ $\vec{x} = A^{-1}T(\vec{x}) + \vec{m}_x$ (recovery)

If approx. then expected mse from $\vec{x}$ & recovered $\vec{x}$ is

$$
\sum_{j=k+1}^{n} \lambda_j .
$$

> Exhibits better rate-distortion tradeoff compared to DFT or the like.

<u>However</u> Computation is costly;

matrix A depends on the input x

Discrete Cosine Transform.

Alg JPEG - DCT

Given an 8×8 image g:

Subtract 128 for every value, every pixel

FOR colors $c = Y, Cb, Cr$ do

FOR frequencies $u, v \in \{0, 1, \dots, 7\}$ do

$$G_{c,u,v} \leftarrow \frac{1}{4}\alpha(u)\alpha(v)\sum_{x=0}^{7}\sum_{y=0}^{7} g_{x,y} \underline{\cos\frac{(2x+1)u\pi}{16}} \cos\frac{(2y+1)v\pi}{16}$$

where $\alpha(u) = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } u = 0 \\ 1 & \text{o/w} \end{cases}$

Make transform orthonormal

end FOR

end FOR

return matrices G_Y, G_{Cb}, G_{Cr}. w/ rounded values

> DCT approximates KL-transform better than DFT in terms of rate-distortion & mse.

$\cos\frac{(2x+1)u\pi}{16} \rightarrow$ rotate $\frac{u\pi}{8}$ for every incr of x

However, if $u=0$, does not rotate

Quantization

Alg JPEG - QUANT

Given an 8×8 ~~imag~~ matrix g with vals $\in [-128, 127]$

Let T be a quantization matrix

$g_{i,j} \leftarrow \frac{g_{i,j}}{T_{i,j}}$ for all $i, j \in \{0, 1, \dots, 8\}$

return g w/ rounded vals

Quantization matrix : 2D DCT basis별 고유 "페널티"

① Basis penalty increases as frequency increases

② Basis penalty for Cr, Cb higher than Y.