N94- 25509

DCT Quantization Matrices Visually Optimized for Individual Images

Andrew B. Watson
NASA Ames Research Center

205749
P. 15

ABSTRACT

This presentation describes how a vision model incorporating contrast sensitivity, contrast masking, and light adaptation is used to design visually optimal quantization matrices for Discrete Cosine Transform image compression. The Discrete Cosine Transform (DCT) underlies several image compression standards (JPEG, MPEG, H.261). The DCT is applied to 8x8 pixel blocks, and the resulting coefficients are quantized by division and rounding. The 8x8 "quantization matrix" of divisors determines the visual quality of the reconstructed image; the design of this matrix is left to the user.

Since each DCT coefficient corresponds to a particular spatial frequency in a particular image region, each quantization error consists of a local increment or decrement in a particular frequency. After adjustments for contrast sensitivity, local light adaptation, and local contrast masking, this coefficient error can be converted to a just-noticeable-difference (jnd). The jnds for different frequencies and image blocks can be pooled to yield a global perceptual error metric. With this metric, we can compute for each image the quantization matrix that minimizes bitrate for a given perceptual error, or perceptual error for a given bitrate.

Implementation of this system demonstrates its advantages over existing techniques. A unique feature of this scheme is that the quantization matrix is optimized for each individual image. This is compatible with the JPEG standard, which requires transmission of the quantization matrix.

DCT Quantization

- Uniform quantization by division and rounding

$$u_{ijk} = \mathrm{Round}\left[c_{ijk}/q_{ij}\right]$$

- Quantization Matrix: 8 x 8 matrix of divisors q_{ij}
- Quantization error

$$e_{ijk} = c_{ijk} - q_{ij}\, u_{ijk}$$

Perceptual Approach to QM Design

- Maximum error is $q_{ij}/2$
- Measure detection thresholds t_{ij}
- Develop comprehensive formula

$$t_{ij} = \mathrm{ap}[i, j, L, px, py]$$

- Set maximum error to detection threshold
- Equivalently, set QM to twice threshold

$$q_{ij} = 2\ t_{ij}$$

Shortcomings

- Image dependence
- Luminance Masking
- Contrast Masking
- Error Pooling
- Quality Metric
- Image-Dependent Perceptual Approach

Luminance Masking

- Thresholds increase with block luminance
- Define Luminance masked thresholds t_{ijk}
- Use comprehensive formula

$$t_{ijk} = \mathrm{ap}[c_{00k}, i, j]$$

- Or use power-law approximation

$$t_{ijk} = t_{ij}\left(c_{00k} / \bar{c}_{00}\right)^{a_T}$$

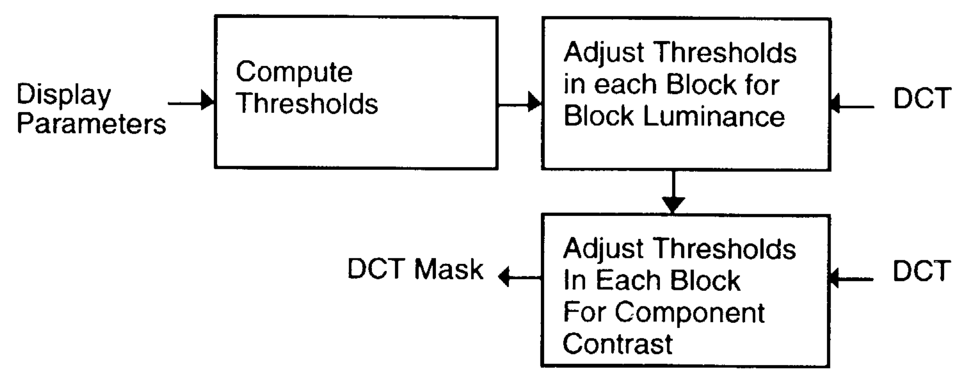
Contrast Masking

- Thresholds increase as contrast increases
- Masking greatest within block and coeff
- Define Contrast-Masked threshold

$$m_{ijk} = \mathrm{Max}\left[t_{ijk}\ , \left|c_{ijk}\right|^{w_{ij}} t_{ijk}^{\ 1-w_{ij}}\right]$$

- w_{ij} (0 - 1) defines strength of masking
- w_{ij} may differ for different frequencies i,j

Computing the DCT Mask



Perceptual Error

- Define elementary perceptual error as *jnd*
- Quantization error divided by masked threshold

$$d_{ijk} = e_{ijk} / m_{ijk}$$

Spatial Error Pooling

- Minkowski metric to pool between blocks

$$p_{ij} = \left(\sum_{k} |d_{ijk}|^{\beta_s} \right)^{1/\beta_s}$$

- Result is "Perceptual Error Matrix"
- Describes the jnds pooled over all blocks at each frequency
- β_s defines nature of spatial pooling

Frequency Error Pooling

- Pool over frequencies to get total perceptual error P
- Minkowski metric

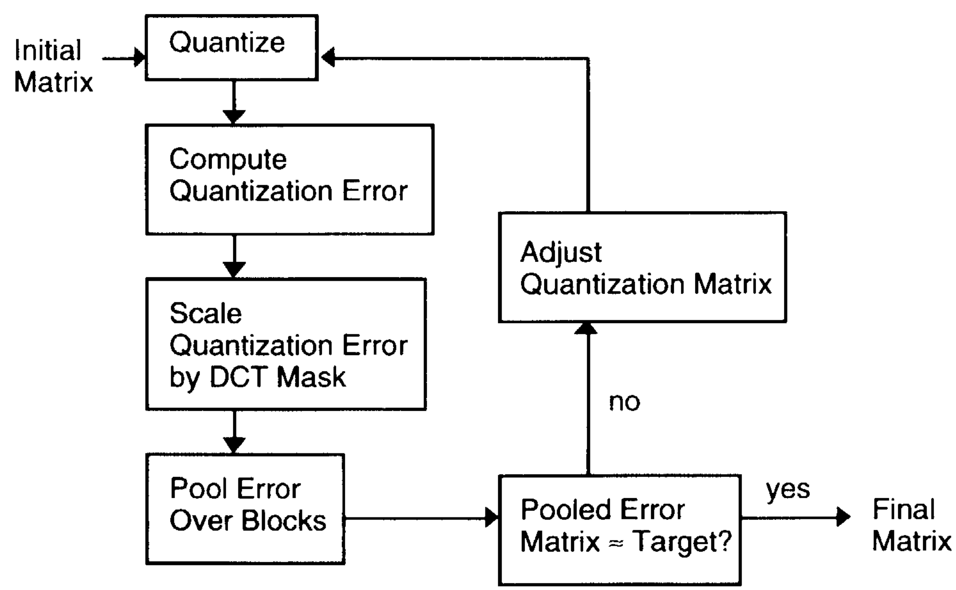
$$P = \left(\sum_{ij} p_{ij}^{\beta_f} \right)^{1/\beta_f}$$

- When $\beta_f = \infty$, is max-of pooling

Optimizing the Quantization Matrix

- Goal: Minimize total perceptual error for given bitrate
- When $\beta_f = \infty$, optimum is when $p_{ij} = \psi \quad \forall \; i,j$
- Intermediate goal: find q_{ij} for which $p_{ij} = \psi \quad \forall \; i,j$
- Note that $p_{ij} = \mathrm{f}_{ij}\left(q_{ij}\right)$

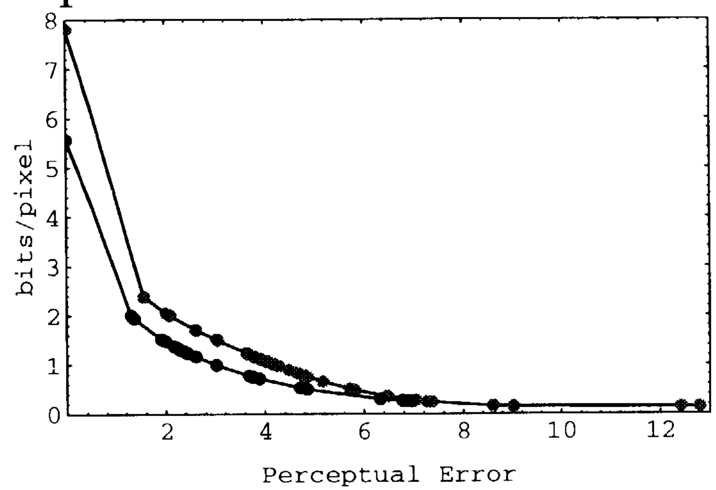
Inner Optimization Block Diagram



C-5

Optimizing for Given Bit Rate

- Samples from function bitrate(ψ)



-
- Iteratively estimate ψ yielding desired bitrate

Summary

- Perceptual error metric based on DCT
- Incorporates luminance masking, contrast masking, and error pooling
- Offers plausible "quality factor"
- Allows simple optimization of QM
- Compatible with JPEG standard
- Can incorporate color & alternate visual models
- Consider the alternatives

Summary (cont.)

- Use in adaptive DCT schemes
  - MPEG
  - thresholding
- Use in wavelet schemes
- "Free"