# Milestones and Trends in Image Compression

*William A. Pearlman*

*Center for Image Processing Research*

*Rensselaer Polytechnic Institute*

VCIP 2010 Huang Shan, PRC 07/12/10

# Outline

- Key historical developments
- Where are we?
- Trends. Improvements?  Breakthroughs?
- Conclusions

Rensselaer

CIPR

# Beginnings of Image Coding

- Shannon 1948:
  - entropy is lowest bit rate possible for perfect recovery
- Shannon 1960:
  - Rate-distortion function (R(D)) gives lowest bit rate possible for reconstruction with distortion no greater than D

# Considerations

- Theorems true for statistically stationary processes – *images not stationary*
- Optimal in limit of long data length
  - not practical
- Early image coding techniques built for stationary models
  - Often simple Gaussian or Laplacian models
  - Parameters may vary block-wise or region-wise

Rensselaer

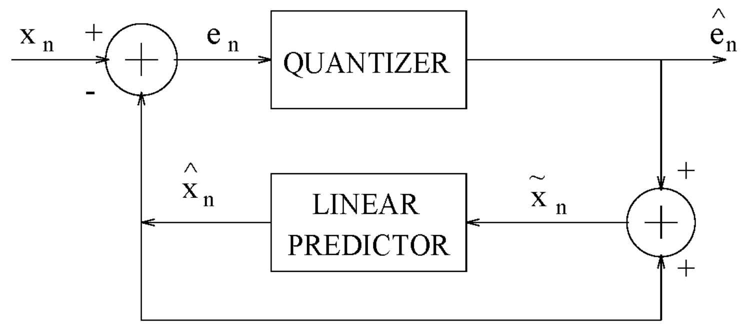
CIPR

# Early Post-Shannon Image Coding

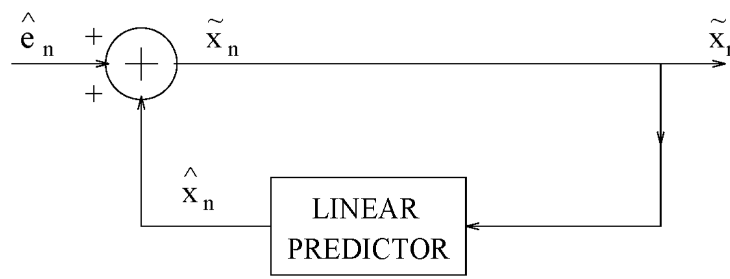
Good reviews: Netravali, Limb, Proc. IEEE 03/80;
Connor, Brainard, Limb, Proc. IEEE 07/72

- DPCM : Prediction and coding from past quantized residuals
  - Quantization tailored to visual perception
  - Predictor from simple to adaptive

MIT (Schreiber, Proc IEEE 3/67),
Bell Labs (Kretzmer, Limb, etc.)




(a) DPCM Encoder



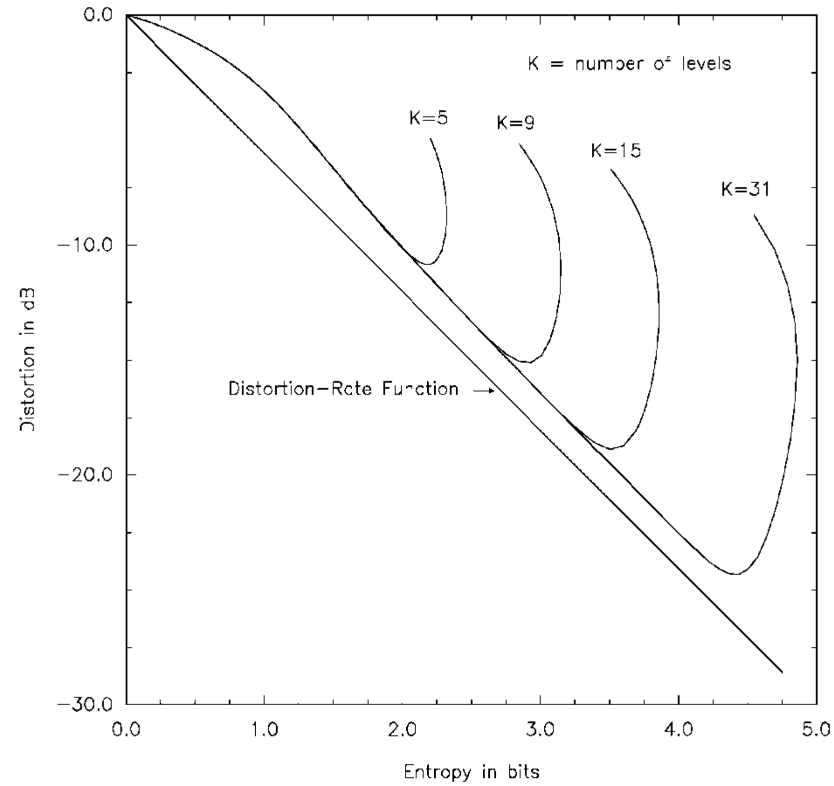

(a) DPCM Decoder

# Advent of Transform Coding

- Huang and Schultheiss (IEEE Trans Commun. Tech. (T-CT) 1963)
  - Proved coding gain for correlated Gaussian sequence via optimal bit allocation to KLT
  - Sparked image transform coding research:
    - Purdue: Habibi and Wintz  (T-CT1971), Wintz (Proc. IEEE 1972)
    - USC: Pratt and Andrews(1968-9), Chen and Pratt (T-COM 1974)
    - MIT: Anderson and Huang (T-CT ’71), Woods and Huang (Picture Bandwidth Comp Wkshp 1969)

# Interesting Developments

- Goblick and Holsinger (T-IT 4/67)
  - Entropy coding of outputs of uniform quantizer nearly optimum: ¼ bit > R(D) for Gaussian, MSE
  - Proved more formally for other statistics by Gish and Pierce (T-IT 9/68)




❖Validates uniform quantizer as choice for minimum MSE with given entropy.

# Transforms

- Search for easily computable, fixed transforms
  - KLT optimal: dependent on statistics and no fast alg.
  - DFT asymptotically optimal, fast algorithm
  - Hadamard fastest to compute, but inefficient in coding
  - Slant, SVC harder to compute, more efficient in coding
- Applied to 16x16 image blocks
- DCT : Ahmed, Natarajan, Rao (T-Cmptr 01/74)
  - Approached KLT spectrum closely for finite N
  - Fixed, independent of statistics, with fast algorithm
  - Became dominant, canonical transform

# Region Adaptive Transform Coding

- Chen and Pratt, “Scene Adaptive Coder”, T-Comm 1984
  - Divided 16x16 DCT blocks into 4 classes, calculated 4 intra-class variance distributions for rate allocation
  - Forerunner of JPEG standard
- JPEG Standard 1989-1992
  - Codes 8x8 DCT blocks independently with Huffman coding of uniform step size quantizer outputs
  - Huffman code based on statistics gathered from experiments with a large number of images

# Vector Quantization

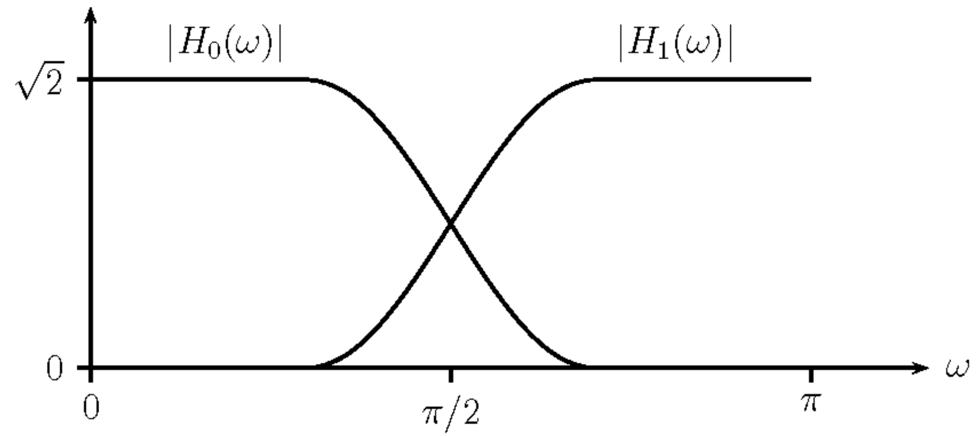
- Generalized Lloyd or LBG (Linde-Buzo-Gray) algorithm (T-COM, 01/80)
  - Asymptotically optimal: complexity ~ $2^{nR}$
  - Restriction to small n and statistical mismatch limited performance
- TCQ (Trellis Coded Quantization) (Marcellin & Fischer,T-COM, 1/90)
  - Asymptotically optimal: complexity ~n
  - Within 0.21 dB of R(D) for Gauss iid source (>1 b/s)
  - Deteriorates in performance < 1 b/s
  - Adopted in JP2000, Part II for Wavelet TCQ

# Subband Coding

- Woods and O'Neil (ICASSP 4/86,T-ASSP 10/86); Gharavi and Tabatabai (VCIP 10/86)
  - First subband coding of images, although done before for speech
  - DPCM coding in subbands
  - Superior results over DCT coding
    - Surprising performance – theoretically unjustified by W&O analysis, eventually justified by P. Rao, S. Rao, and Pearlman (T-IT 3/91, 7/96)

# Wavelets

- Need alias-cancelling half-band filters (a low-pass and a high-pass) for perfect reconstruction
- QMF and paraunitary filters were exact or approximate solutions



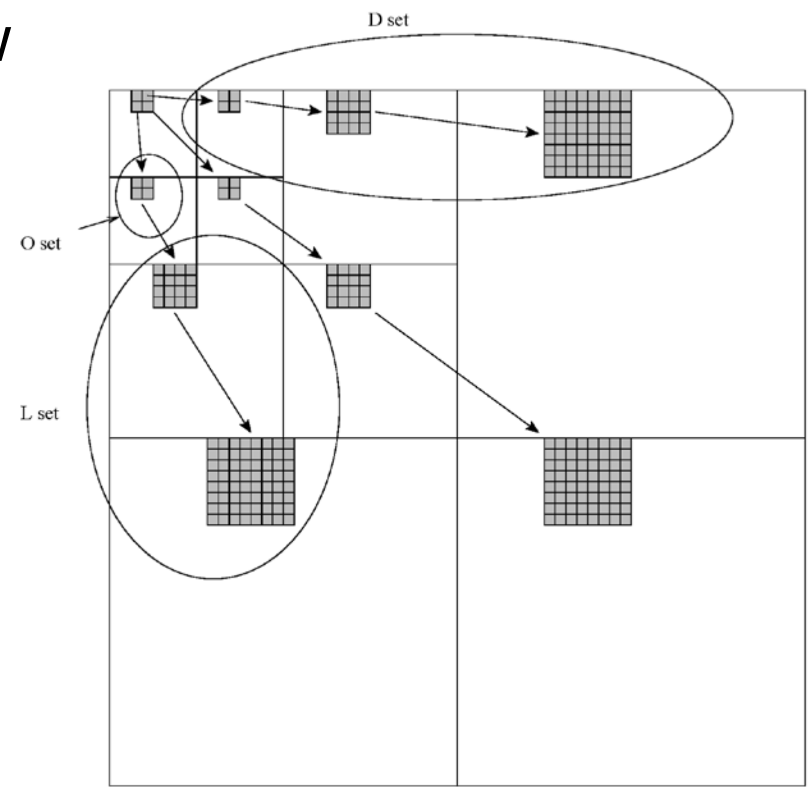

- Then came wavelet filters, specifically the CDF 9/7 biorthogonal wavelet filter
  --- used first for image coding by Antonini, Barlaud *et al.* (T-IP 4/92)

# Coding of Subbands

- Kinds of coding of subbands independently
  - Huffman coding, arithmetic coding, predictive coding (DPCM), tree coding, trellis coding, VQ, TCQ
    - All did well, the more complex the better the result
- Significant breakthrough: *zerotree coding*
  - EZW (Shapiro, ICASSP'92, T-ASSP 12/93)
    - Takes advantage of decaying amplitude with wavelet subband frequency

# SPIHT

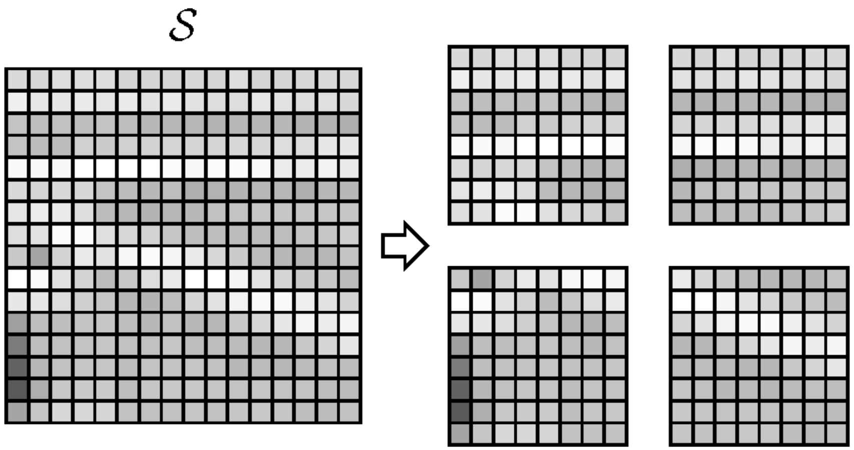
- SPIHT (Said and Pearlman, T-CSVT 6/96)
  - Introduces set partitioning in spatial orientation trees with roots in lowest frequency subband
  - Finds groups of pixels below set of thresholds $T=2^n$
    -- reduces to n raw bits for smallest n (+ sign bit)
    -- n=0: 1 ‘0’ bit locates group




❖Amplitude-based, non-statistical bit assignment
❖Simple arithmetic operations
❖More efficient than EZW

# Another Partition (SPECK)

- SPECK (Islam and Pearlman, VCIP99)
  - Recursive quadri-section of blocks
  - Quadtree code for execution path



# JPEG2000

- Seek better low rate performance than JPEG
  - Eliminate blocking artifacts found in low-rate JPEG
    - LOT (Malvar, 1992) was one solution
- Embed new features: rate scalability, ROI encoding/decoding, etc. (inherent in SPIHT and EZW)
- Codes subblocks of wavelet subbands with EBCOT coder (Taubman, T-IP 07/00)
  - Subblocks' size 64x64 or 32x32
  - Context-based adaptive arithmetic bitplane coding
- Part 1 finalized in 2001

# Trends

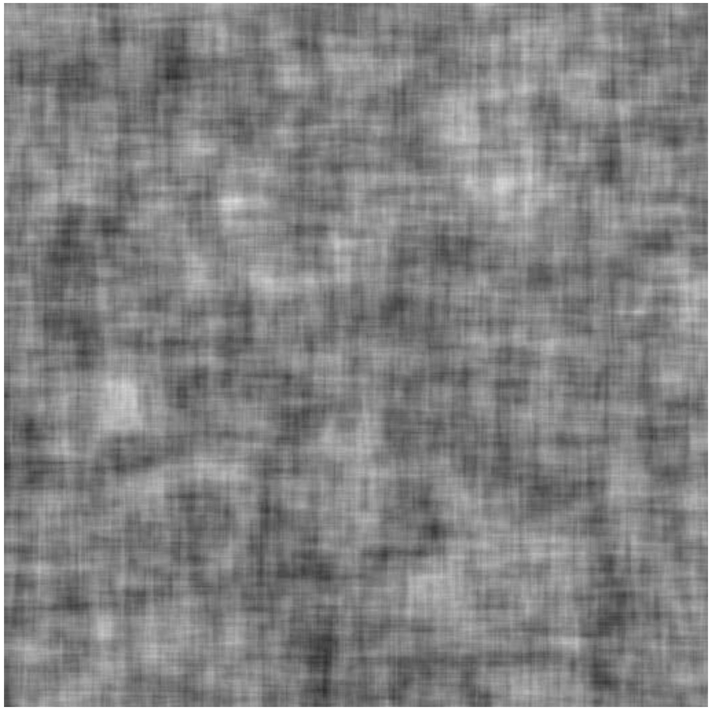
- Get closer to code based on actual value
  - SPIHT does it almost perfectly
- Smaller coding units
  - Enables finer resolution, locally adaptive coding
    - 8x8 DCT in JPEG, 4x4 DCT in H.264/AVC, JPEG XR
    - Subband subblock coding in JPEG2000
- Overlapped blocks or inter-block prediction
  - Eliminates discontinuities at block boundaries
- More complex, adaptive context-based entropy coding (e.g., JPEG2000, H.264/AVC)
- Simpler block transforms (integerized DCT)
  - Less decorrelation compensated by more complex coding

# Efficiency of Modern Methods

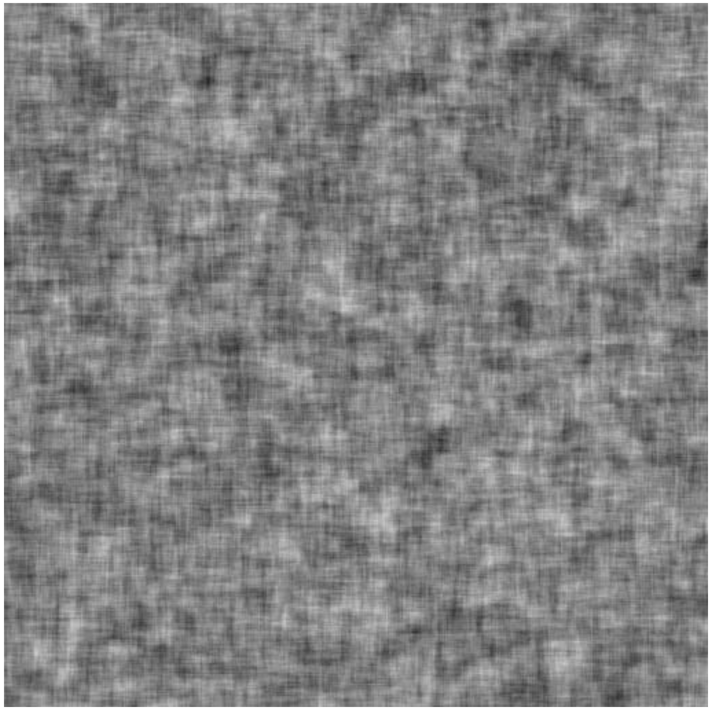
- Methodology
  - Generate Gauss-Markov Images
  - Compare compression results with Rate-Distortion or joint entropy function

# Gauss-Markov Images

Variance = 400    Mean = 128



a = 0.95

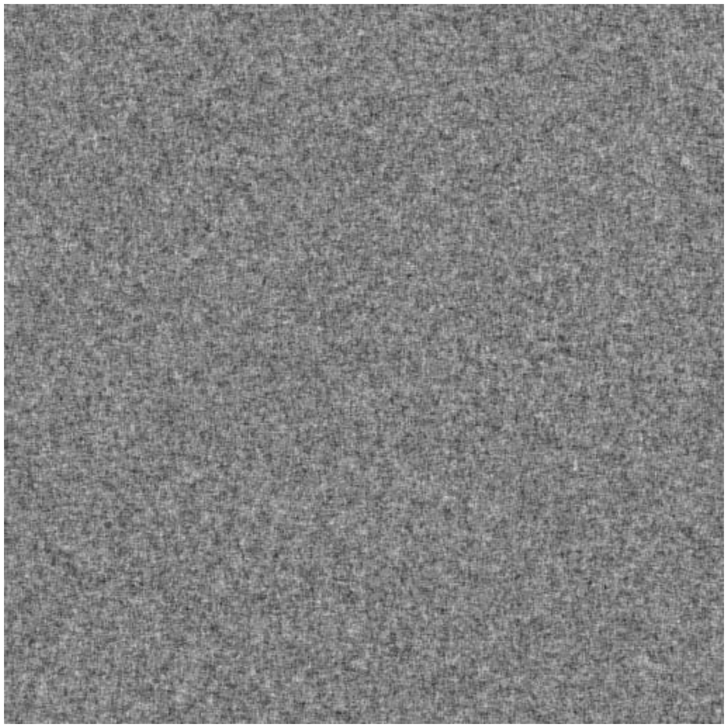


a = 0.90

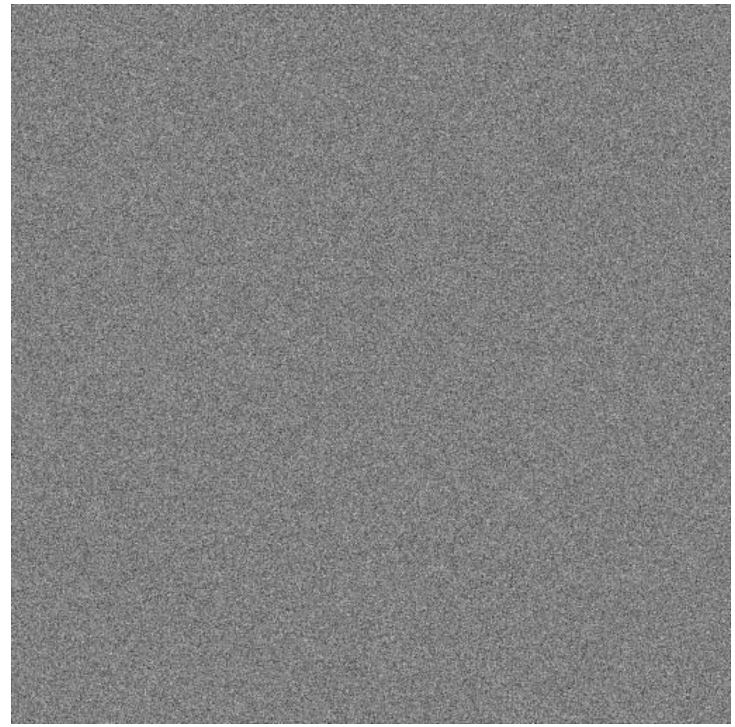
Separable;  8-bit precision;
512x512 lower cut from 640x640

# Gauss-Markov Images (cont.)

Variance = 400 Mean = 128



a = 0.50



a = 0.0

Separable; 8-bit precision;
512x512 lower cut from 640x640

# Theoretical Bounds

Rate-Distortion Function (Gaussian, squared error)

$$R = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \max\{0, \tfrac{1}{2} \log_2 \frac{\lambda(i)\lambda(j)}{\theta}\}$$

$$D = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \min\{\theta, \lambda(i)\lambda(j)\}$$

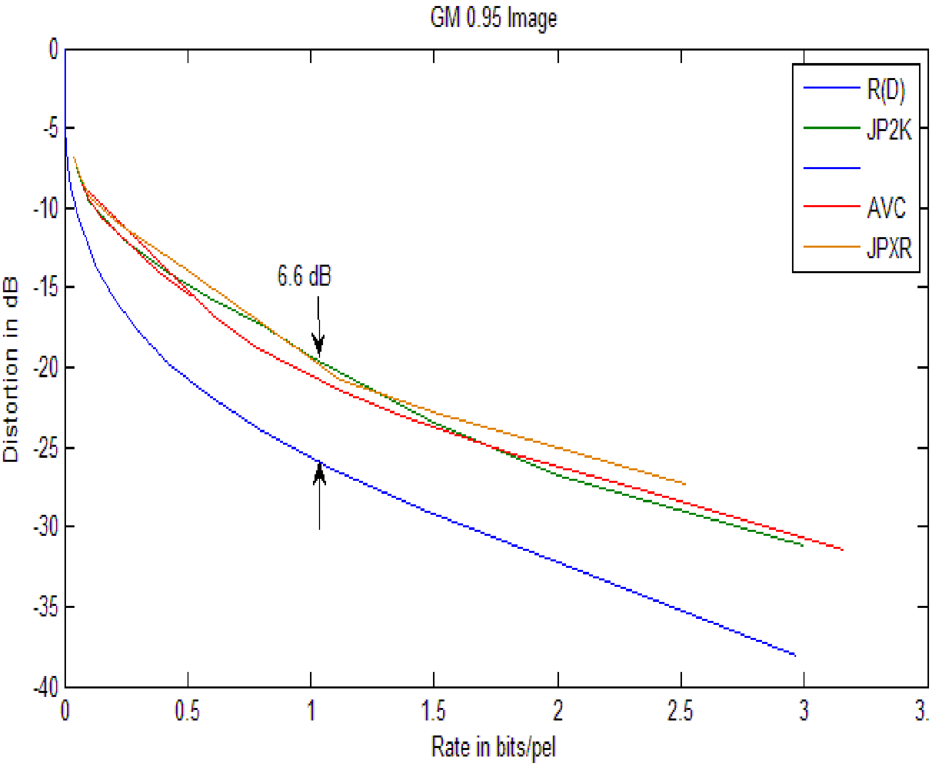
Approximate Gaussian Entropy Function

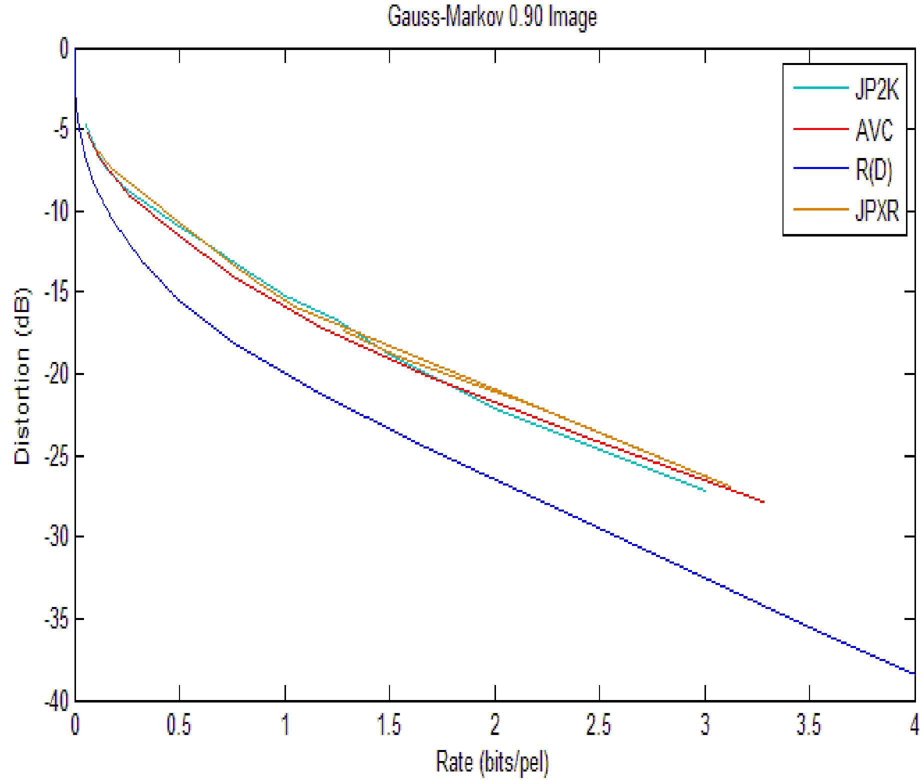
Using $H(X^n) \approx h(X^n) - \log_2 \delta$ $\quad -\log_2 \delta = m$ bits (precision)

$$\frac{1}{N^2} H(X^{N^2}) \approx \frac{1}{2N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \log_2(\lambda(i)\lambda(j)) + \tfrac{1}{2} \log_2 2\pi e \sigma_X^2 + \frac{m}{N^2}$$

(Eigenvalues normalized for unit variance)

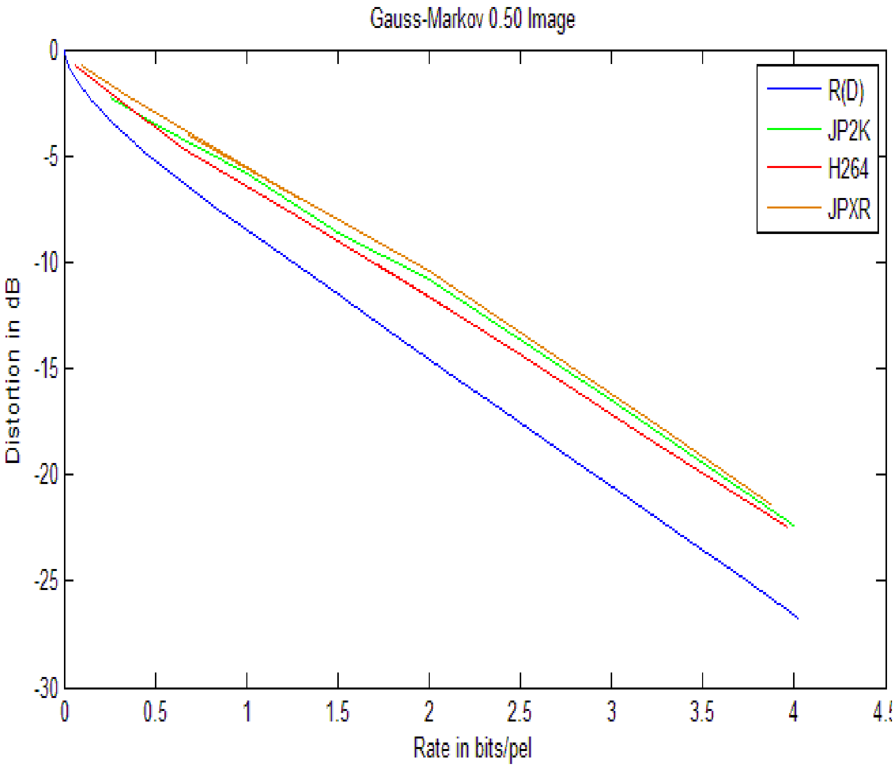
# Comparisons

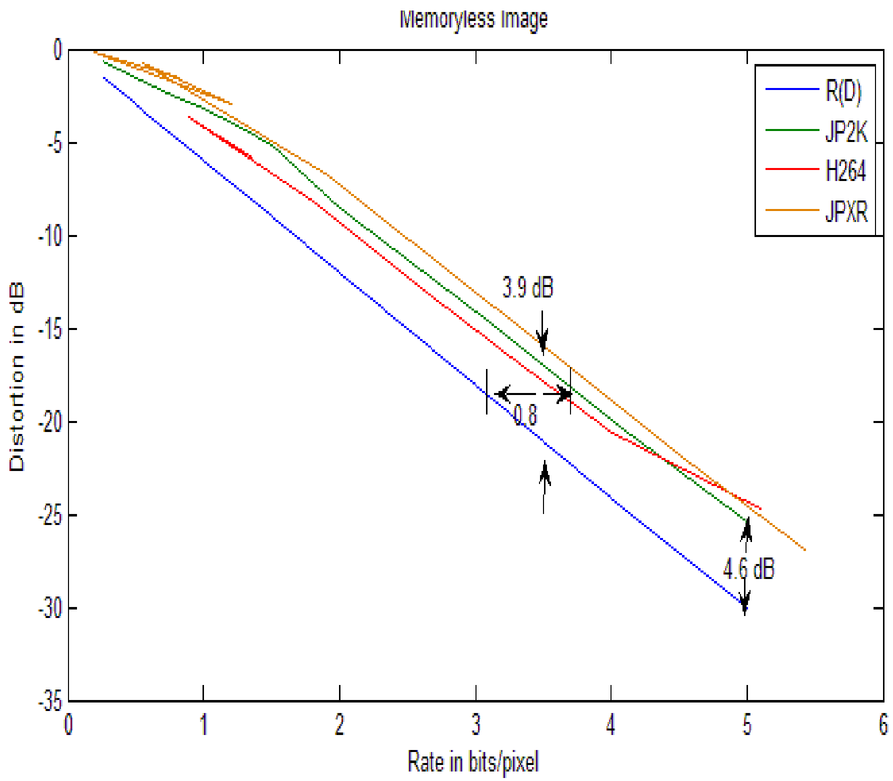



PSNR = 22.11–D dB

# More Comparisons




PSNR = 22.11–D dB

Rensselaer

CIPR

# Lossless Compression

| | | Differences from Entropy (b/p) | | | |
|---|---|---|---|---|---|
| Correlation Parameter | Joint Entropy | SPIHT | CALIC | JP2K | JPEG-XR |
| 0.95 | 3.0172 | 0.9438 | 0.4188 | 0.4838 | 0.9329 |
| 0.90 | 3.9778 | 0.6652 | 0.2312 | 0.3762 | 0.6847 |
| 0.50 | 5.9548 | 0.2392 | 0.0872 | 0.3172 | 0.2939 |
| 0 | 6.3691 | 0.2469 | 0.1779 | 0.3639 | 0.4420 |

* CALIC closest to entropy in all cases
* Aside from CALIC, SPIHT at a = 0.5 and 0 beats others

# What Have We Learned?

- Much room for improvement for lossy compression :
  - > 0.5 bpp for high quality
  - 4 to 6 dB at useful bit rates
- Small room for improvement for lossless compression - ~0.2 bpp

**Lesson: The best adaptive techniques can take you only so far.

# Where to go from here?

- For pure compression, much more potential payoff for lossy methods.
- Clearly advantageous to transform to independent variables and/or segment to stationary entities.
  - closes performance to the latter gaps
- Barring advancements in pure compression, need to pursue
  - better transforms that are adaptive to image features
    - Bandelets, curvelets, etc. ?
  - better segmentation and set partitioning methods

# Future Application Space

Question: *Do we need more efficient compression?*

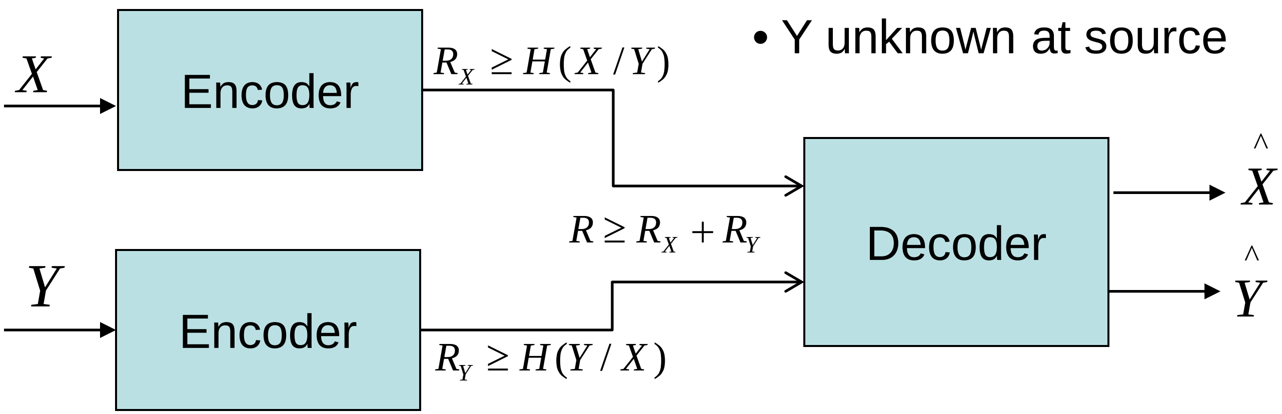
- Large images with multiple dimensions
  - Examples:
    - 4 dimensions: fMRI, medical ultrasound
    - Materials micro-structures with many attributes at given grid point.
- Content-based retrieval from large databases
  - Internet application needs interactivity for consultation and quantitative analysis.
  - Need fast search and retrieval and fast scalable decoding for browsing, retrieval, and transmission
    - Places limits on complexity and memory usage
      - Increase in size always seems to outpace gains in speed
    - Not likely to close existing performance gaps with simpler techniques that utilize less memory.
      - Fruitful or fruitless pursuit?
    - Contribution is to limit degradation the least possible by being clever

❖ Transmission rate decrease is main motivation for compression
  - Storage reduction now secondary

# Distributed Source Coding

Source Coding with Side Information: Slepian-Wolf 1973, Wyner-Ziv 1976

- X and Y correlated sources
- Y unknown at source




S-W: Encode X with H(X/Y) bits, Y with H(Y) bits, can achieve $\hat{X} = X$

*No loss over when Y is known at encoder also, if statistics X given Y are known.*

W-Z: *Lossy coding performance same whether Y is known at both ends or only at decoder, if statistics of X and Y are jointly Gaussian.*

# DSC Image Compression Scenarios

- Low complexity encoding for image transmission
- Sensor networks
  - Multiview coding
- Multiple description coding
- Camera alignment
- Cryptogram compression

❖ None likely to bridge identified performance gaps, especially for the usual non-Gaussian lossy coding

# Technology Advances

- Dramatic increases in processor speeds seem to be ending
  - Parallelization by multi-core processor chips is the trend
    - Power consumption $\sim f^3$
  - New parallel forms of algorithms for compression likely to emerge
    - Currently JPEG2000, JPEG, etc. have parallel structure --- currently not exploited
    - Multiple description coding; distributed source coding
- More compact, higher power batteries would expand application scenarios for compression
- Miniaturization to quantum limit to be reached in 10 to 15 years
  - Quantum Computers: lower rate limits theoretically possible

# Quantum Computing

- Quantum computers can solve some math problems considerably faster than classical computers
- Qbit.com (defunct) – claimed 2-10:1 lossless image compression at 1.5 Gbits/sec throughput
  - with qubit processor?  US 2004/0086038 App.
- Quantum Information Theory
  - Well developed; parallels Shannon theory
    - Source coding theorem (von Neumann entropy limit)
    - R(D) theorem
    - S-W and W-Z theorems
    - Channel capacity theorem
  - Theoretically achievable rates lower than in classical computing

# Quantum Bits and Entanglement

- General state of one qubit (input): $\alpha$ 's complex

$$|\psi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle \,,\; |\alpha_0|^2 = \Pr\{|0\rangle\} \quad |\alpha_1|^2 = \Pr\{|1\rangle\} \;,\quad |\alpha_0|^2 + |\alpha_1|^2 = 1$$

  - said to be *entangled*

  Ex.: photon $|\psi\rangle = (1/\sqrt{2})|0\rangle + (1/\sqrt{2})|1\rangle$ linear polarized at 45°

- Output is measurement: $|0\rangle$ or $|1\rangle$
  - Orthogonal states can be measured
  - Similarly for 2-qubit system- states are *entangled*

$$|\psi\rangle = \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle$$

- *n-qubit space – $2^n$ dimensional Hilbert Space*
- *States can not be copied or cloned.*
- *A measurement changes the state:* basis of secure key distribution
- *States can be communicated*

# Entropy Example

Two equiprobable photon states:   Shannon entropy = 1 bit

Suppose 0 → $|0\rangle$    H polarization

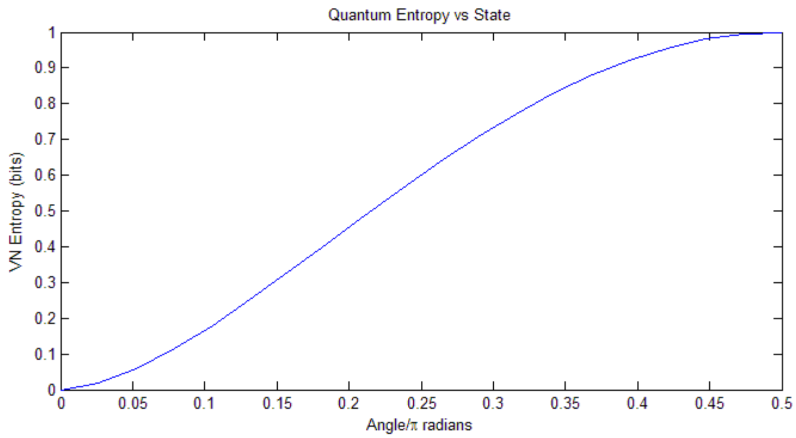
Suppose 1 → $\psi = \cos\theta|0\rangle + \sin\theta|1\rangle$    Angle $\theta$ polarization

Von Neumann Entropy S($|0\rangle$, $\psi$ ) = $_\theta$

$H_2((1-\cos\theta)/2)$

$H_2(p) = -p\log_2 p - (1-p)\log_2(1-p)$

(Binary entropy function)




Except for $\theta = \pm\pi/2$ ,    S($|0\rangle$,$\psi$ ) < 1   (e.g., 0.60 at θ = π/4)

But, only $\theta = \pm\pi/2$ is detectable or communicable!!

❖ Therefore, von Neumann entropy not yet physically realizable.

# Prospect of Lower Compression Limit

- So far, quantum information theory does not give physically realizable lower entropy limits
- Also, the devices and detectors work only in the laboratory or with limited capability – polarizers,1-qubit gates, and short shift registers
- Short error-correcting codes, secure key distribution
- Physicists are hard at work to make the devices that form and detect specified quantum states
- Physicists have taken the lead at formulating quantum information theory, but IT community has been roused (e.g., Devetak & Berger, "Quantum R-D Theory," Trans. IT Jun 2002; Rob Calderbank)
- Further reading
  - E. Desurvire, *Classical and Quantum Information Theory (Cambridge 2009)*
  - M. A. Nielson, I. L. Chang: *Quantum Computation and Quantum Information*
  - N. D. Mermin : *Quantum Computer Science: An Introduction*
  - J. Audretsch, Ed.: *Entangled World: The Fascination of Quantum Information and Computation*
  - Bennett & Shor, "Quantum Information Theory", Trans IT, Oct 1998

# Conclusion

- Substantial gaps to compression limits still exist
- Trend toward algorithms working in small coding units and using complex entropy coding
- Trend to multiple core processors to spur development of new parallel processing paradigms
  - Collaborative compression
- Open question whether quantum information theory and quantum computation will bring future rate savings

*Thank you!*