The Need for Better Models for Coding Sparse Multimedia Representations

Amir Said

Hewlett Packard Labs, Palo Alto

July 2005

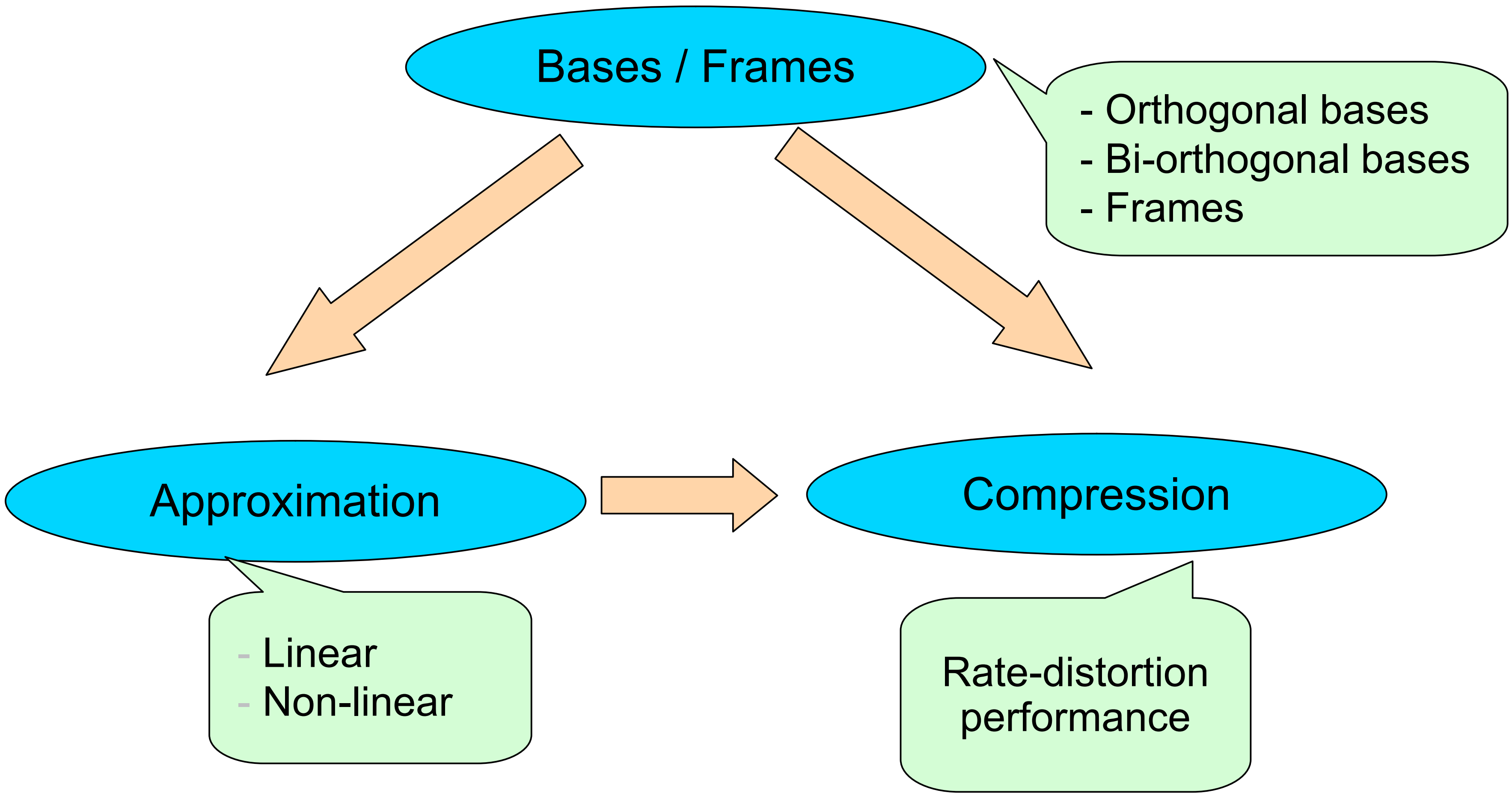
Multimedia Compression

- **Very high economic significance**
- **Difficult to evaluate theoretical limits**
  - Some believe much better compression is possible
- **Research field somewhat idiosyncratic**
  - Some “standards committee” mentality
    - Only the best “wins”!
    - After choosing the “winner” we are “done”
  - Some impatience with scientific process
    - Multimedia compression research is repeatedly considered “dead”
  - Empiric and ad hoc techniques can predominate

Multimedia Signals

- Commonly represented using linear transforms
  - "Classical": Fourier, Karhunen-Loève, Wavelets, …
  - New construction of bases for natural signals
- Common theme: sparse representation of audiovisual information is key in
  - Signal analysis (detection, classification, recognition)
  - Signal representation (compression)
  - Signal enhancement (denoising)

Basic Concepts and Relations



*We know human response needs to be considered too…

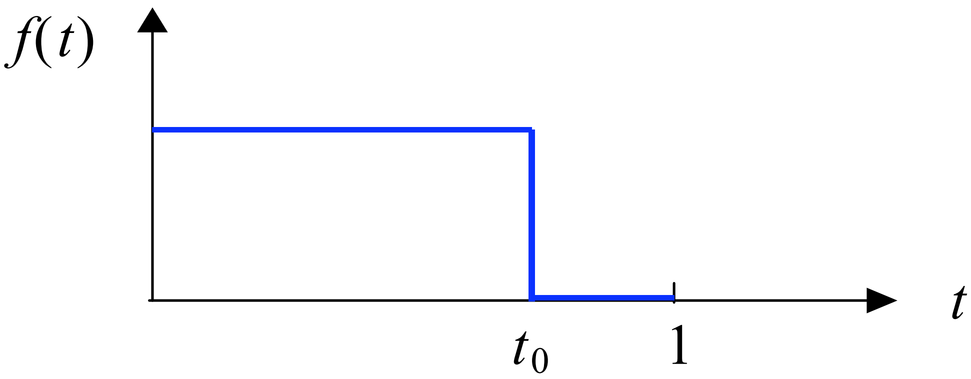
Main Approximation Types

- **Basis functions** $\{\varphi_0, \varphi_1, \varphi_2, \dots\}$
- **Linear approximation (LA)**
  $$\hat{x}_M = \sum_{i=1}^{M} \alpha_i \varphi_i$$
  - Keep first M components
  - KLT is best basis only in very limited sense
- **Nonlinear (adaptive) approximation (NLA)**
  $$\tilde{x}_M = \sum_{i \in I_M} \alpha_i \varphi_i$$
  - Keep best M components
  - Power of nonlinear approximation depends on the basis
- **Approximation quality**
  $$\hat{\varepsilon}_M = \left\| x - \hat{x}_M \right\|_2$$
  $$\tilde{\varepsilon}_M = \left\| x - \tilde{x}_M \right\|_2$$

Approximation Theory

Example:

[Vetterli]



$t \in [0,1]$, t_0 ~ uniform distribution over [0, 1]

	Linear approximation	Nonlinear approximation
Fourier series	$\hat{\varepsilon}_M \sim 1/M$	$\tilde{\varepsilon}_M \sim 1/M$
Wavelet series (Haar wavelets)	$\hat{\varepsilon}_M \sim 1/M$	$\tilde{\varepsilon}_M \sim 1/2^M$

Some Recent Approaches

- **Geometric**
  - Directional transforms (fixed)
    - Ridgelets, curvelets, contourlets, etc.
  - Adaptive geometry-based approaches
    - Beamlets, wedgelets, bandelets, directionlets, etc.
- **Frames and overcomplete representation**
  - Matching pursuits
  - Wavelet packets
  - A variety of best basis searches
  - Several types of adaptive transforms

Compression

- **Approximation quality (distortion)**
- **Description complexity (rate)**
  - LA: code coefficient values only
  - NLA: code both coefficient values & coefficient locations
- **Transform coding**
  - Simple coding may be (much) more effective in the transform domain
  - Entropy coding is quite powerful and general, but severely limited
    - Complexity, time for "learning" statistics, etc.

Coding Sparse Representations

- **Cohen, Daubechies, Guleryuz & Orchard**
  - **"On the importance of combining wavelet-based non-linear approximations with coding strategies," *IEEE TIP*, July 2002.**
- **Realistic rate-distortion analysis is essential**
- **If it is truly sparse then the rate is dominated by bits for locating nonzero coefficients**
  - **Possible to switch coding methods when representation is not sufficiently sparse**
  - **Only at high rates the performance can be significantly limited by scalar quantization**

Coding Values and Location of Transform Coefficients

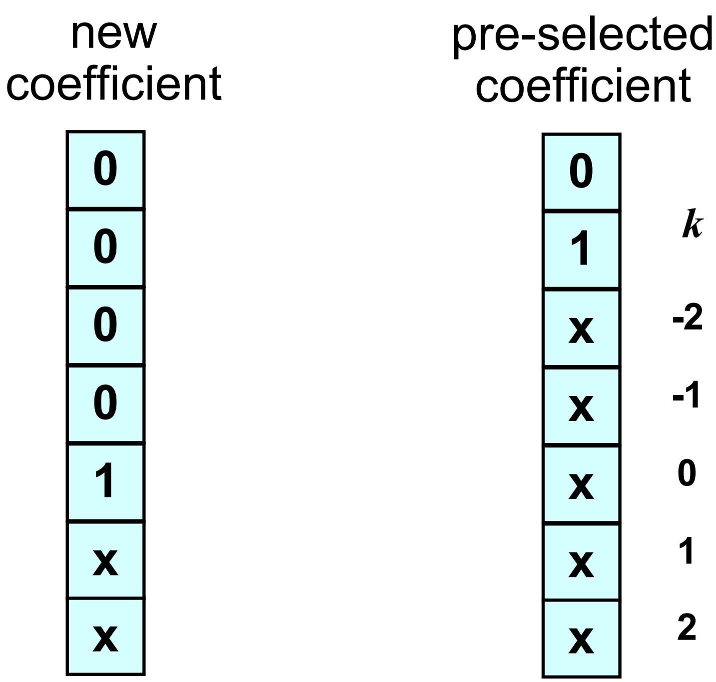
- **Statistics after symmetric quantization**
  - Uniform & deadzone (different interval for zero symbol) quantizers

$$p_0 >> p_1 = p_{-1} >> p_2 = p_{-2} >> \cdots$$

- **Majority of bits used to code, directly or indirectly, location of symbol '0'**
  - Compression performance defined mostly by bits used to code location of symbol '0', '+1', '–1'
- **Advantages of bit-plane coding**
  - Always coding '0', '+1', '–1' symbols, defined by sequence of quantization steps 2^{-k}
  - In a certain way, always the same problem

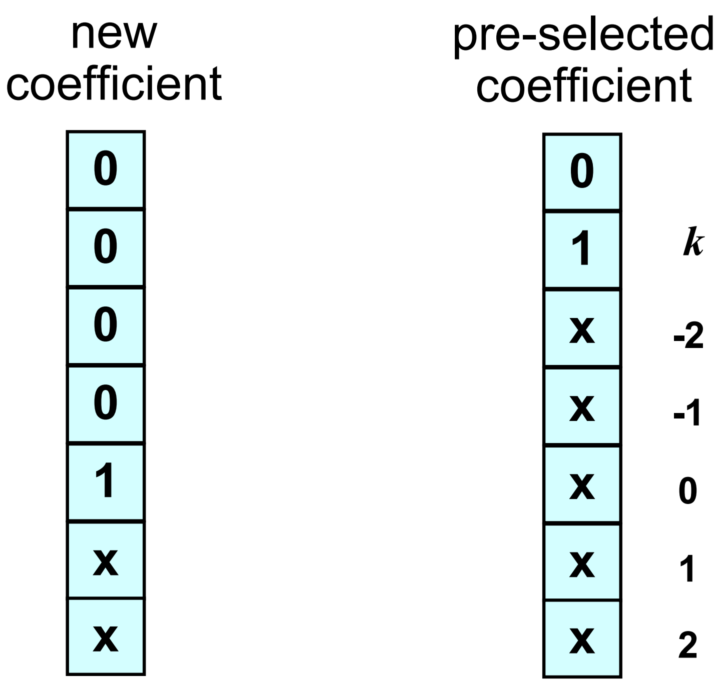
Rate-distortion Analysis for Bit Allocation

- **When to refine or code new coefficients?**



Rate-distortion Analysis for Bit Allocation

- **When to refine or code new coefficients?**



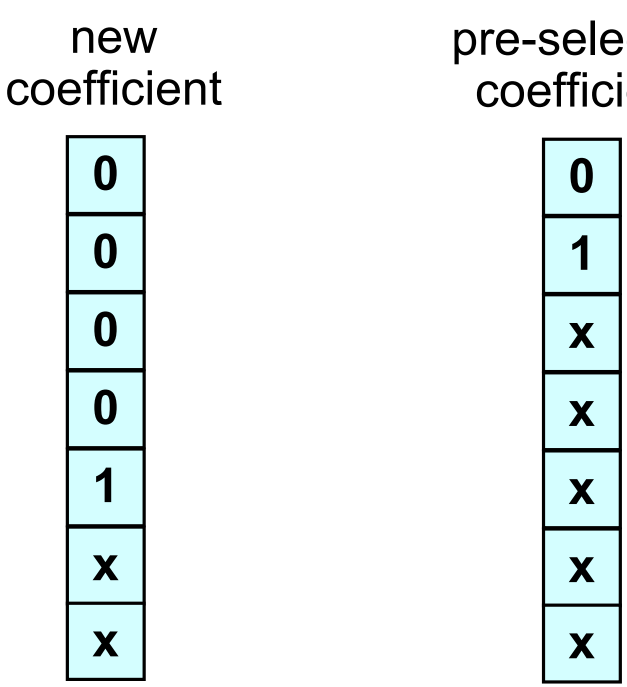
N_{pb} = average number of bits to code new location

$$\frac{\Delta D_N(T)}{\Delta R} \approx \frac{-9T^2}{4(N_{pb}+1)} \qquad \frac{\Delta D_P(T)}{\Delta R} \approx \frac{-T^2}{2^{k+2}}$$

Average distortion change per bit

Rate-distortion Analysis for Bit Allocation

- **When to refine or code new coefficients?**



k	
-2	$N_{pb} < 1.75$
-1	$N_{pb} < 3.5$
0	$N_{pb} < 8$
1	$N_{pb} < 17$
2	$N_{pb} < 35$

Better to code coefficient position whenever

$$N_{pb} < 9 \cdot 2^{k} - 1$$

Entropy Coding

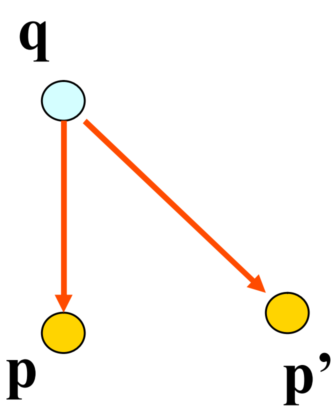
- **Coefficient bits: typically little compression possible**
  - Can be shown to be true in a wide set of conditions using the analysis of a coding technique called symbol grouping (Said, DCC'05)

- **Coefficient location: must be coded exploiting statistical properties of the transform**
  - Good models can yield better compression…

Redundancy Analysis

$$\ell(\Psi,\mathbf{p}) = \sum_{n=1}^{N} \sum_{s \in G_n} p_s \log_2\left(\frac{p_s}{\bar{p}_n}\right)$$

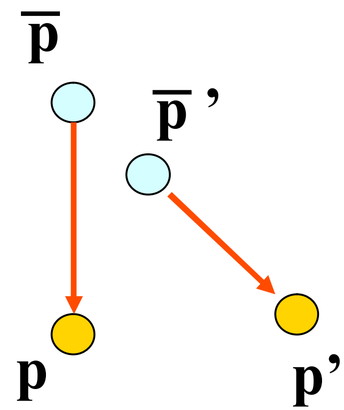
Similar to Kullback-Leibler distance



Code mismatch redundancy

$$D(\mathbf{p},\mathbf{q}) = 0 \quad \Leftrightarrow \quad \mathbf{p} = \mathbf{q}$$

$$\frac{\partial D(\mathbf{p},\mathbf{q})}{\partial p_i} = \log_2\left(\frac{p_i}{q_i}\right) + \frac{1}{\ln(2)}$$



Bit-plane coding redundancy

$$\ell(\Psi,\mathbf{p}) = 0 \quad \Leftrightarrow \quad \mathbf{p} = \bar{\mathbf{p}}$$

$$\frac{\partial \ell(\Psi,\mathbf{p})}{\partial p_i} = \log_2\left(\frac{p_i}{\bar{p}_i}\right)$$

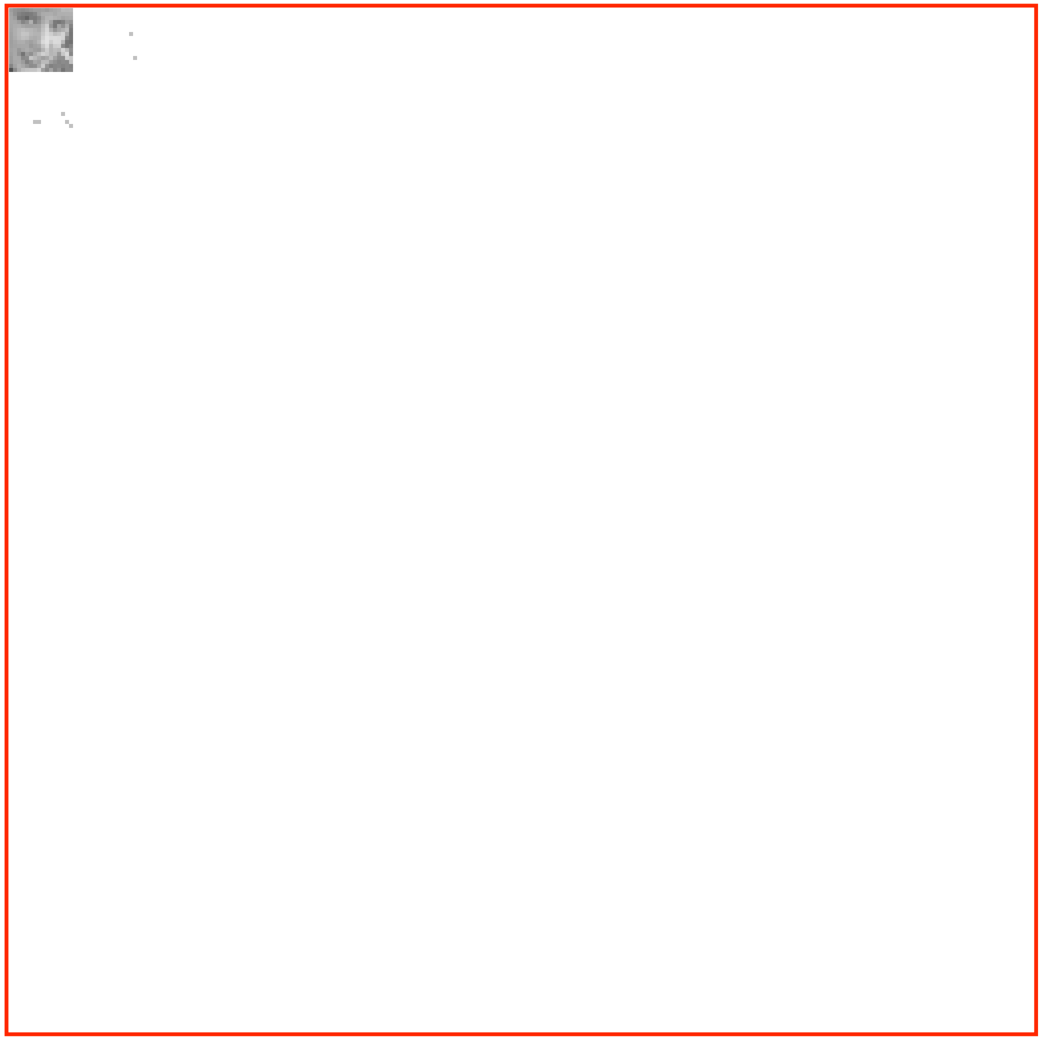
Wavelet Coding Example

Original image

Significance on Bit-plane 9

White = insignificant

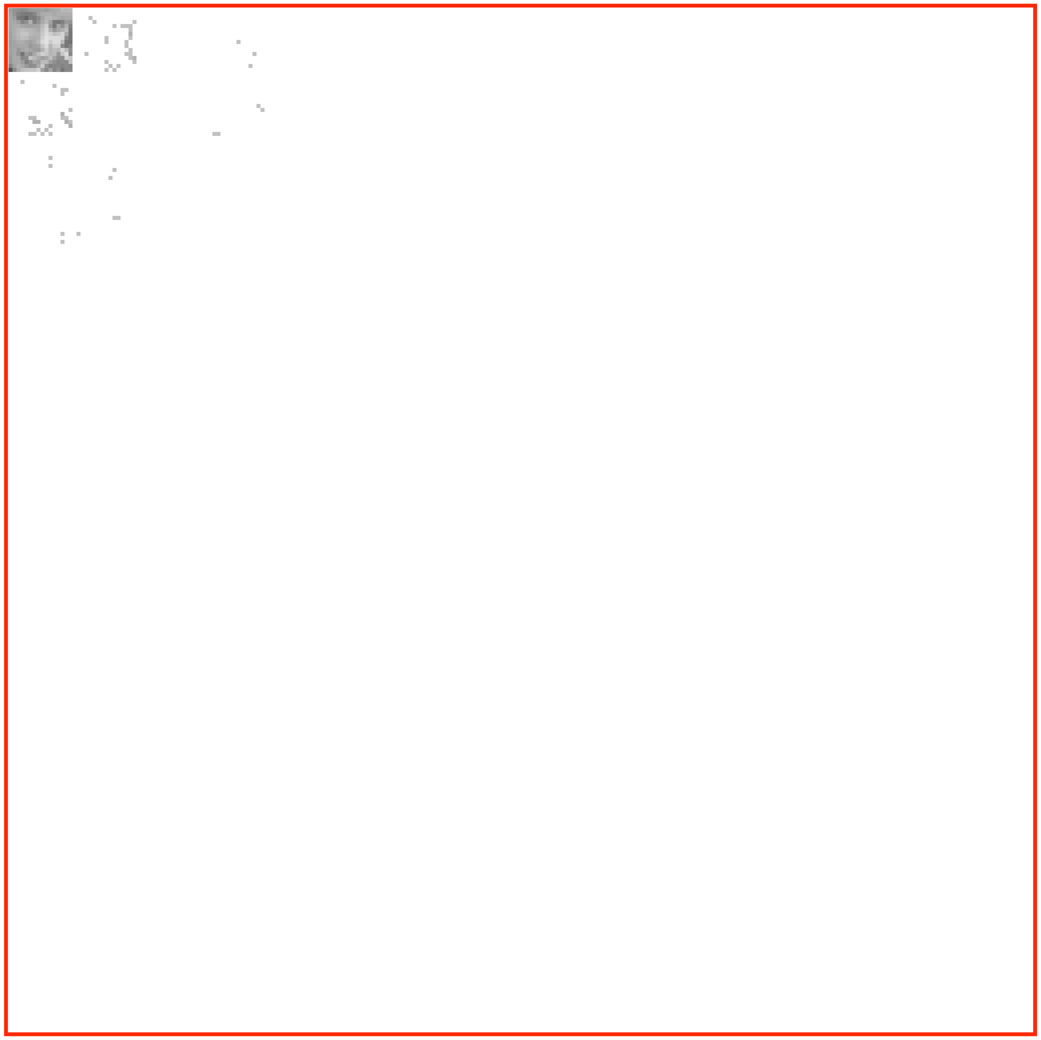
Gray = significant



Significance on Bit-plane 8

White = insignificant

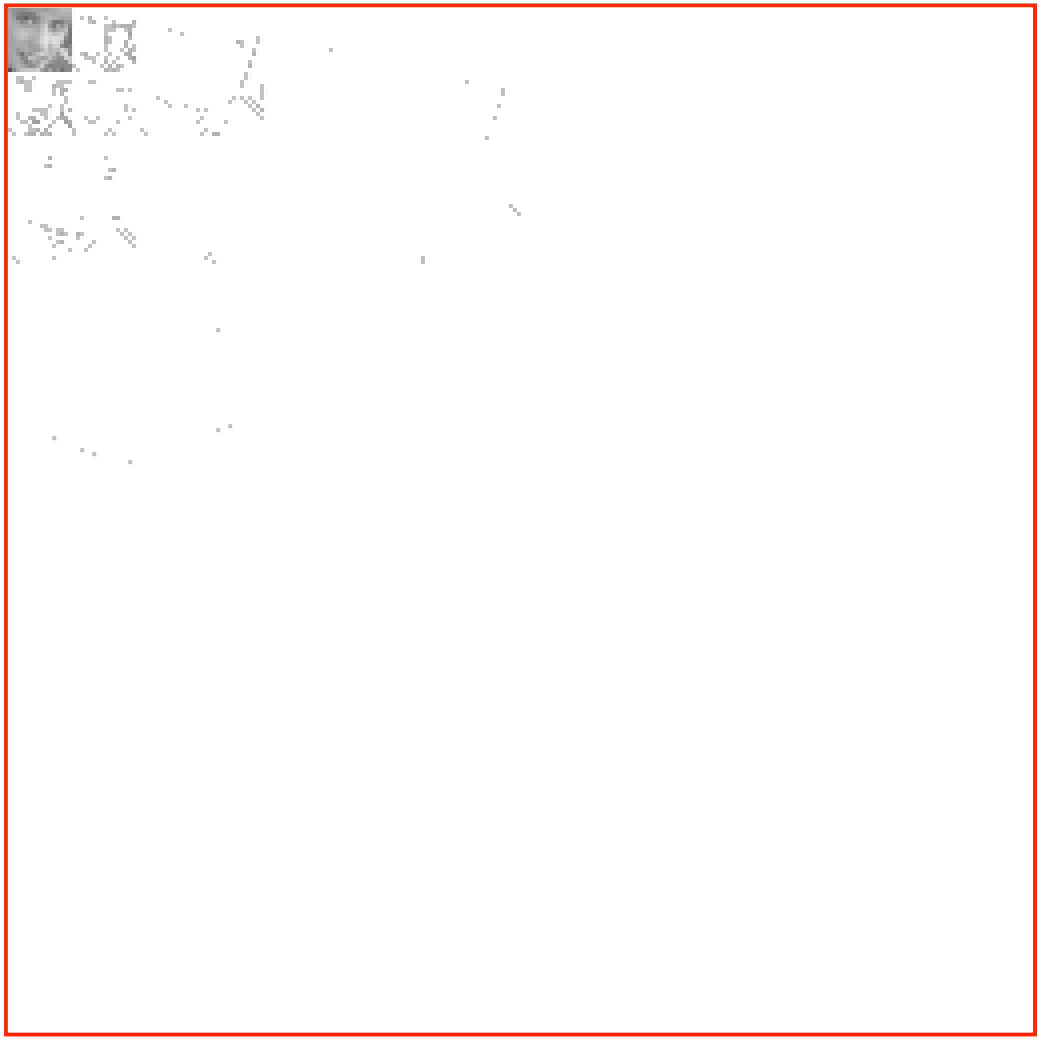
Gray = significant



Significance on Bit-plane 7

White = insignificant

Gray = significant



Significance on Bit-plane 6

White = insignificant

Gray = significant

Significance on Bit-plane 5

White = insignificant

Gray = significant

Significance on Bit-plane 4

White = insignificant

Gray = significant

Significance on Bit-plane 3

White = insignificant

Gray = significant

Significance on Bit-plane 2

White = insignificant

Gray = significant

Significance on Bit-plane 1

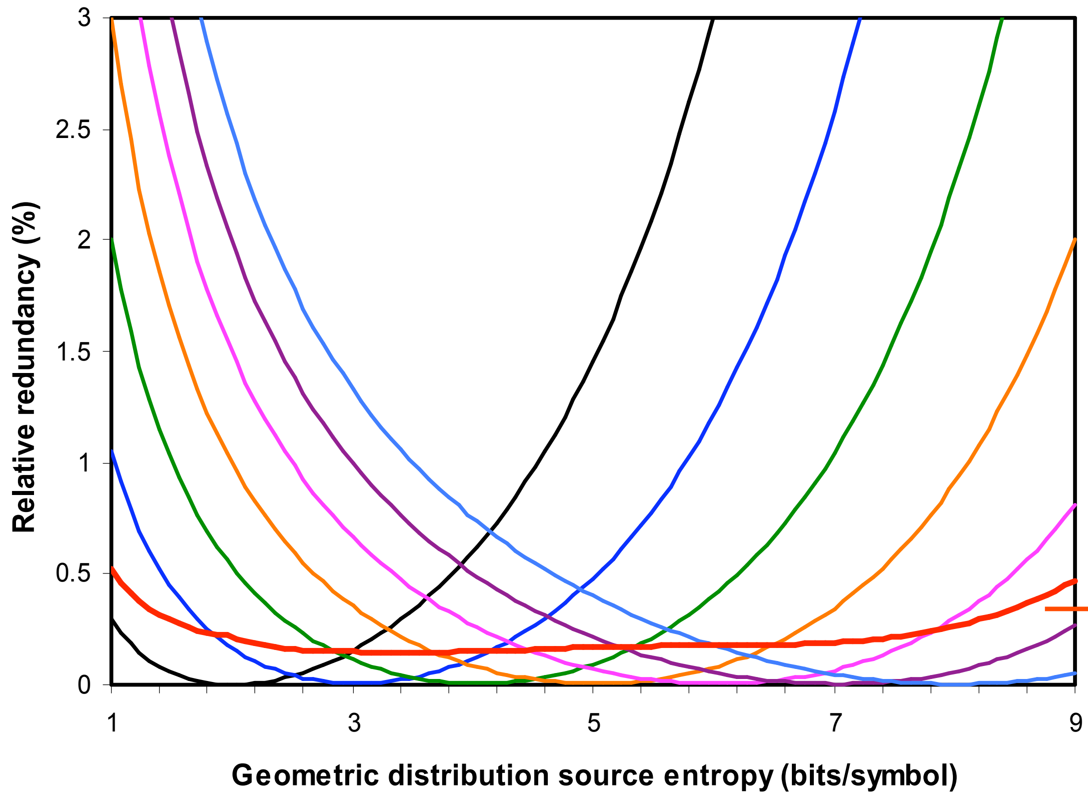
White = insignificant

Gray = significant

Set-partitioning Coding

- **Entropy coding technique for implicitly coding coefficient positions**
  - Can exploit dependence between subbands
- **Bits used for coding coefficient location by SPIHT**
  - Typically start with average of 2.5-3 bits/pixel
  - Increasing to about 9 bits/pixel in higher bit planes, and about 5 bits/pixel in lower bit planes
  - Partially ordered lists are used for near-optimal bit allocation
- **Very similar performance when applied to square blocks inside wavelet bands (like JPEG2000)**
  - SPECK, SBHP, MC-EZBC: shows that set-partitioning compression efficiency not defined only by multiresolution structure

Set Partitioning Coding Characteristics

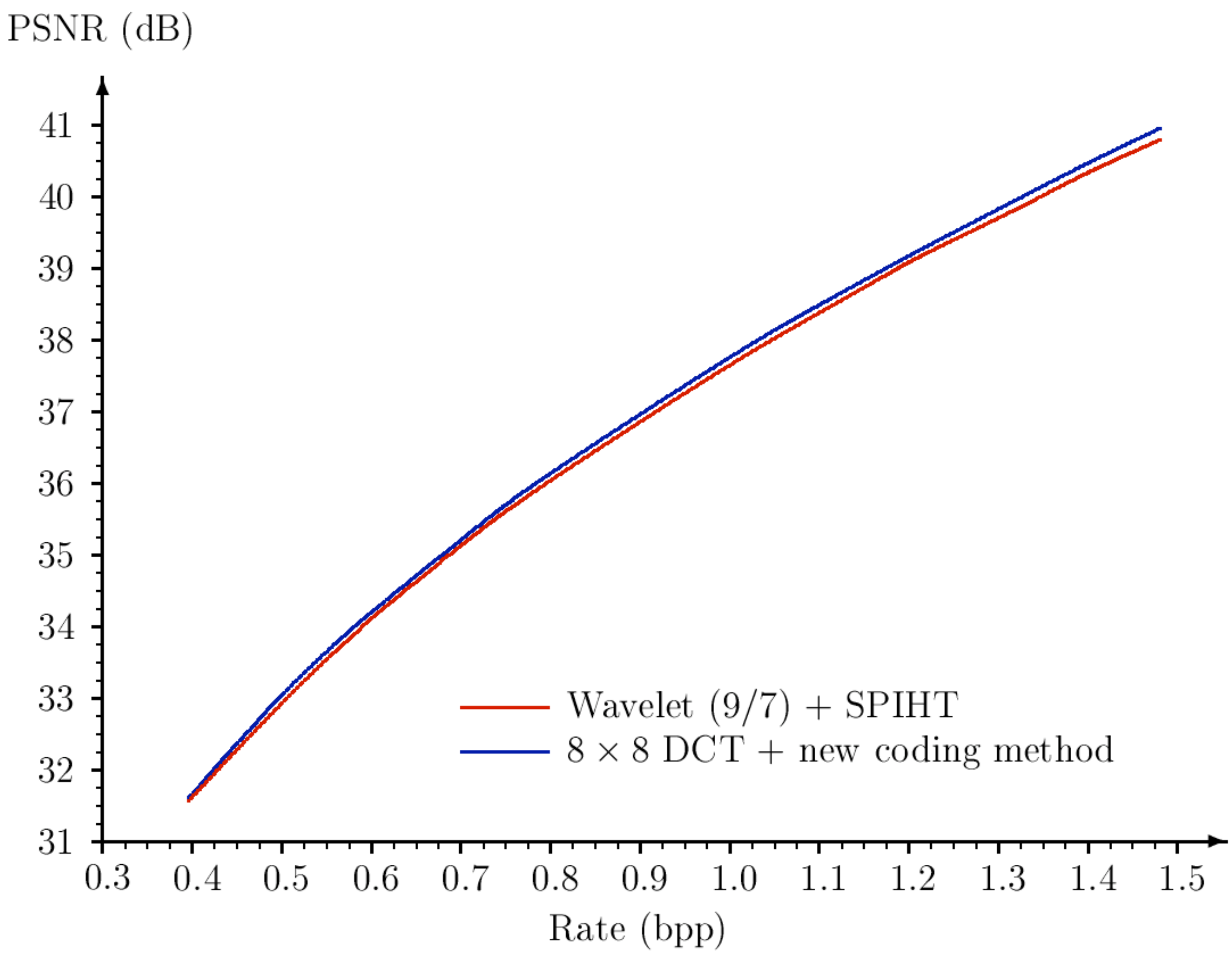


Set partitioning (no arithmetic coding)

All other curves: Golomb-Rice codes

Some New Results

- **ISO image 'Bike'**



New DCT Coding Technique

- **Not using set-partitioning**
  - 0.5-1.0 dB better than results obtained by Xiong, Guleryuz & Orchard, *IEEE SPL*, Nov. 96.
- **Follows several "threads" in the 8×8 block, instead of one zig-zag path**
  - Objective is to more reliably sort probabilities
- **Hundreds of adaptive coding models**
  - They learn efficient image statistical model
  - Limited amount of information for designers…

Other Recent Developments

- Scalable video compression
  - Great advances on the performance of wavelet-based video encoders
  - Still, AVC produced better results…
- AVC intraframe compression
  - directional prediction + block transform
  - competitive with wavelet-based JPEG2000

Conclusions and Suggestions for Discussion

- **Great amount of activity in the development of new sparse representations, nonlinear approximation theory**
- **Usefulness for compression depends on proper coding of nonzero coefficient positions**
- **Coding can only be optimized if we have good models**
- **Are we making real progress?**
  - **E.g., can improvement in denoising lead to better compression?**
  - **Any indication of possible breakthroughs?**

invent