Vector Quantizers for Reduced Bit-Rate Coding of Correlated Sources

Russell M. Mersereau

Center for Signal and Image Processing
Georgia Institute of Technology

Outline

- Cache vector quantization
  - Lossless and lossy cache VQ
  - Replenishment algorithms
  - A hierarchical video coder
- Dynamic codebook reordering
  - Principle
  - Performance on Markov sources
  - Performance on a low bit-rate speech coder

References

- **K.K. Truong** and **R. M. Mersereau**, “Vector quantization video encoder using hierarchical cache memory scheme,” U.S. Patent 5,444,489, August 22, 1995.

- **F.G.B. De Natale**, **S. Fioravanti**,   **D.D. Giusto**, “DCRVQ: a new strategy for efficient entropy coding of vector-quantized images”, *IEEE Transactions on Communications*, Vol. 44, pp. 696 --706, June 1996.

- **G. Shen** and **M.L. Liou**, "An efficient post-processing technique and a window based fast serach algorithm for image vector quantization" IEEE Trans. on CASVT, vol 10, no.6 Sept. 2000 .

- **V. Krishnan**, *A Framework for low bit-rate speech coding in noisy Environments*, Ph.D. Thesis, Georgia Tech ,2005 (**D. Anderson and T. Barnwell** , Advisors).

Cache Vector Quantization (1992)

- A CVQ consists of two (or more) codebooks
  - ✓A small dynamic codebook
  - ✓A large main codebook
- The small codebook contains recently used codevectors
- The small codebook is searched first. If a fit is found the index in the small codebook is sent, else the large codebook is searched.
- The cache is updated after each transmission.

Lossless and Lossy Caches

- Lossless: Use cache entry only with perfect fit.

- Lossy: Accept cache entry if the fit is “good enough”

Quantization Error < Threshold

Adjusting the Threshold

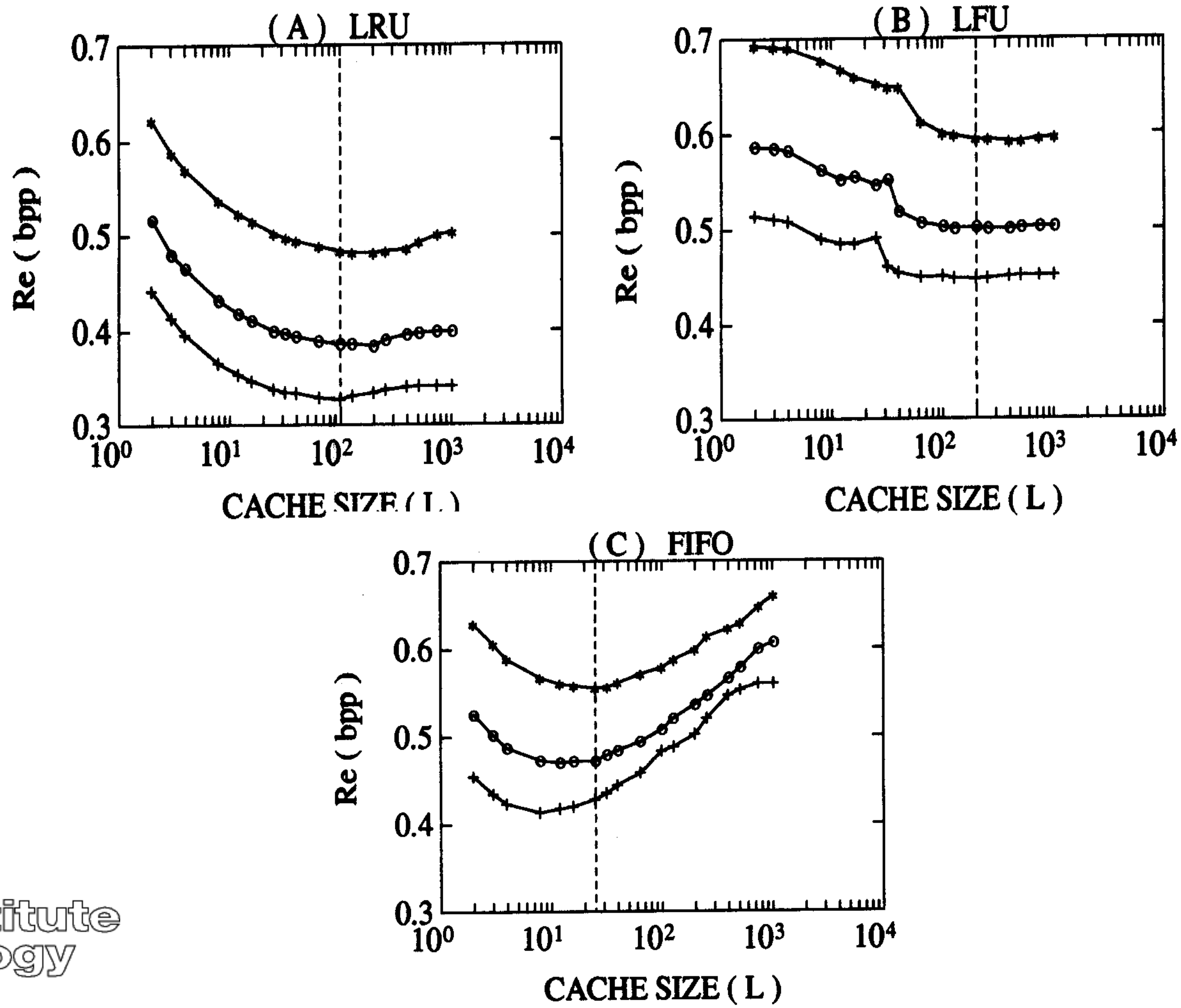
- Raising the threshold, T
  - Speeds up the coder
  - Increases the distortion
  - Decreases the bit rate
- If successive vectors are strongly correlated, CVQ saves both bits and search time.

Cache Replenishment

- A variety of algorithms can be used for cache replenishment
  - FIFO
  - Least frequently used (LFU)
  - Least recently used (LRU)
- For a LRU or LFU cache, the indexes should be entropy coded.

Replenishment Comparison

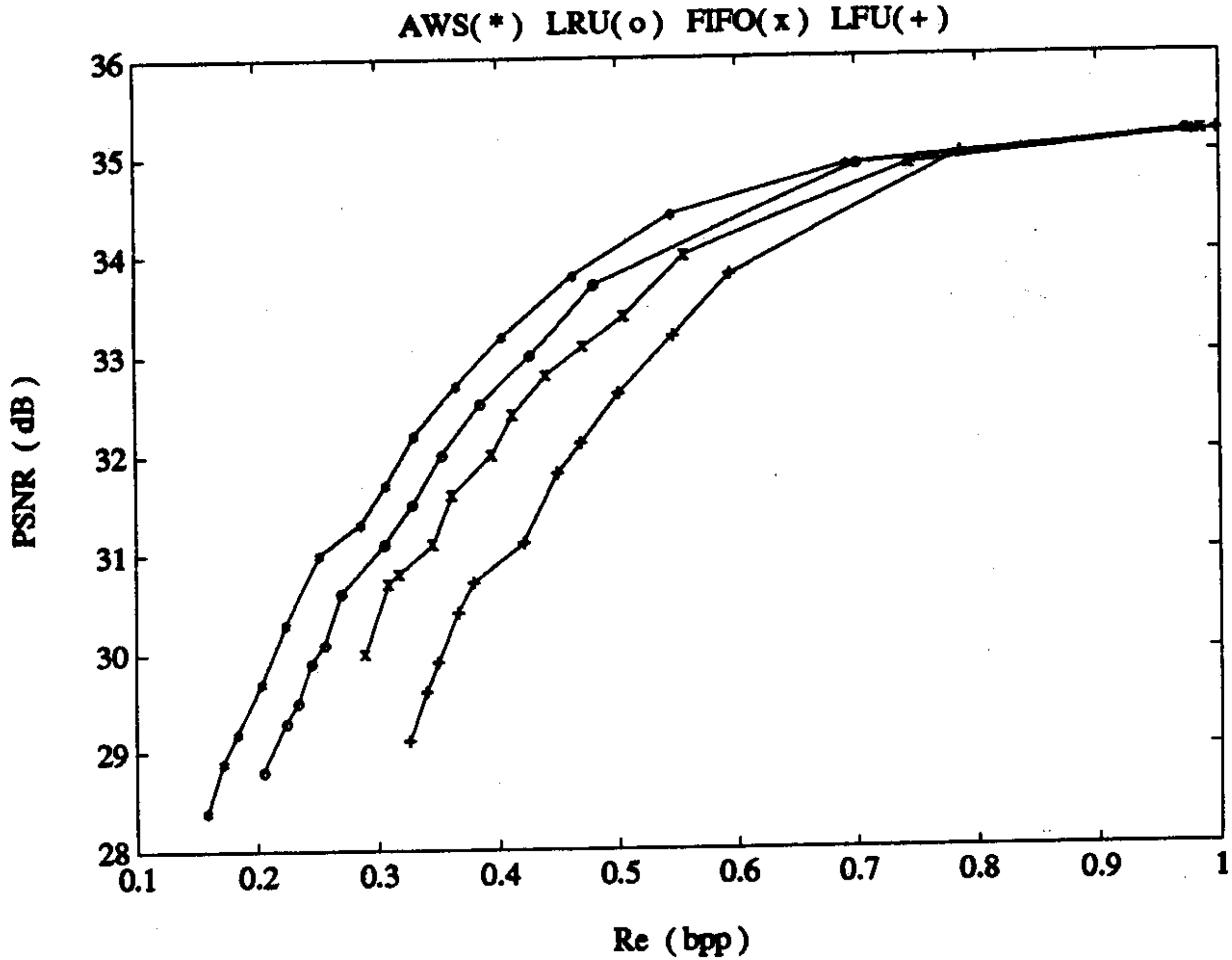
- Coding 4x4 image blocks (Lenna)



Georgia Institute of Technology

CSIP
Center for Signal & Image Processing

Rate-Distortion Performance



Cache Statistics

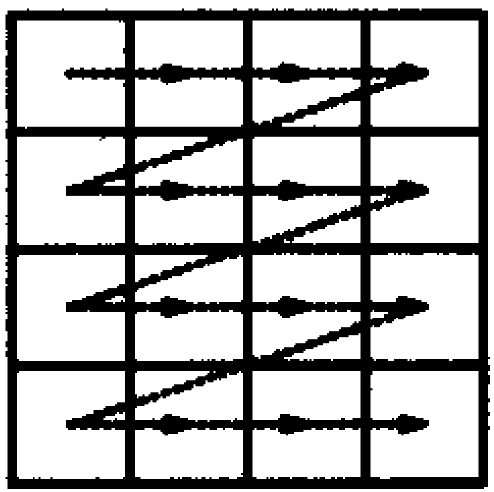
- LRU cache at γ=20

White=cache CB
Black=main CB

Order of Presentation

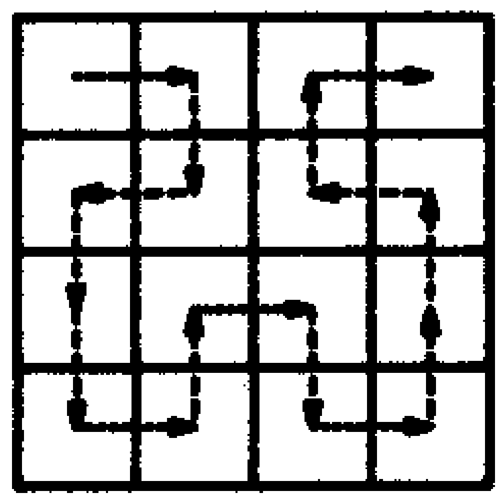
- The performance of the cache depends upon the order of presentation

(A)



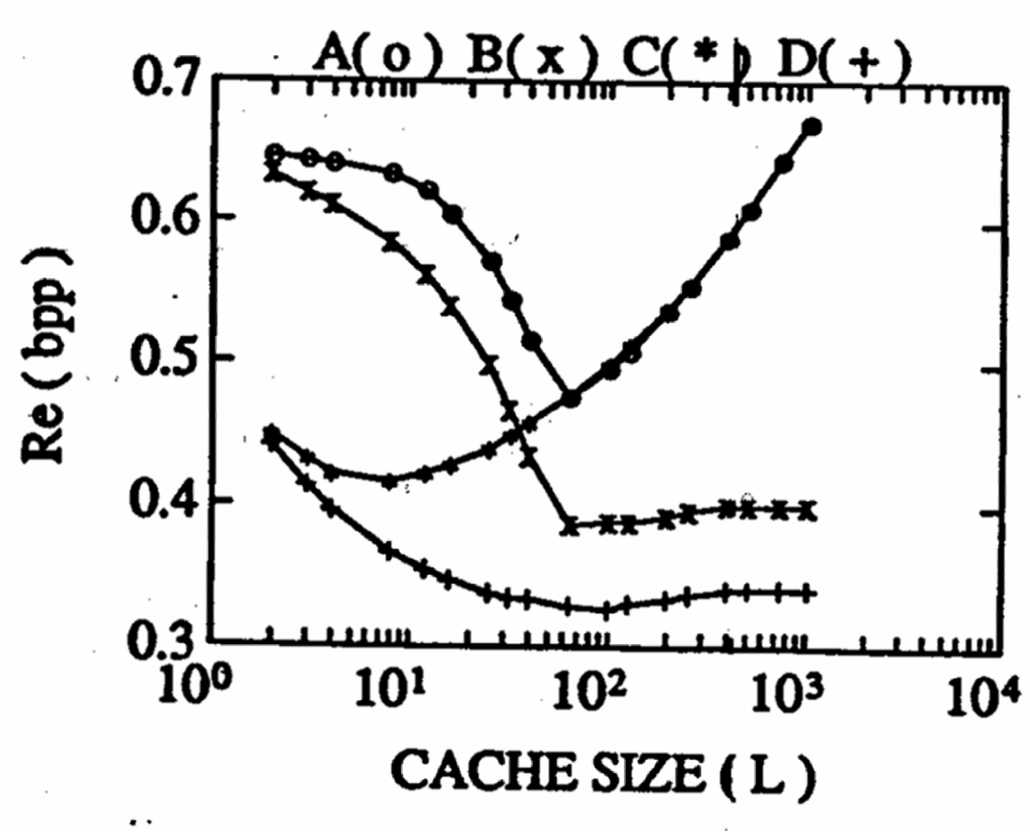
Raster scan

(B)



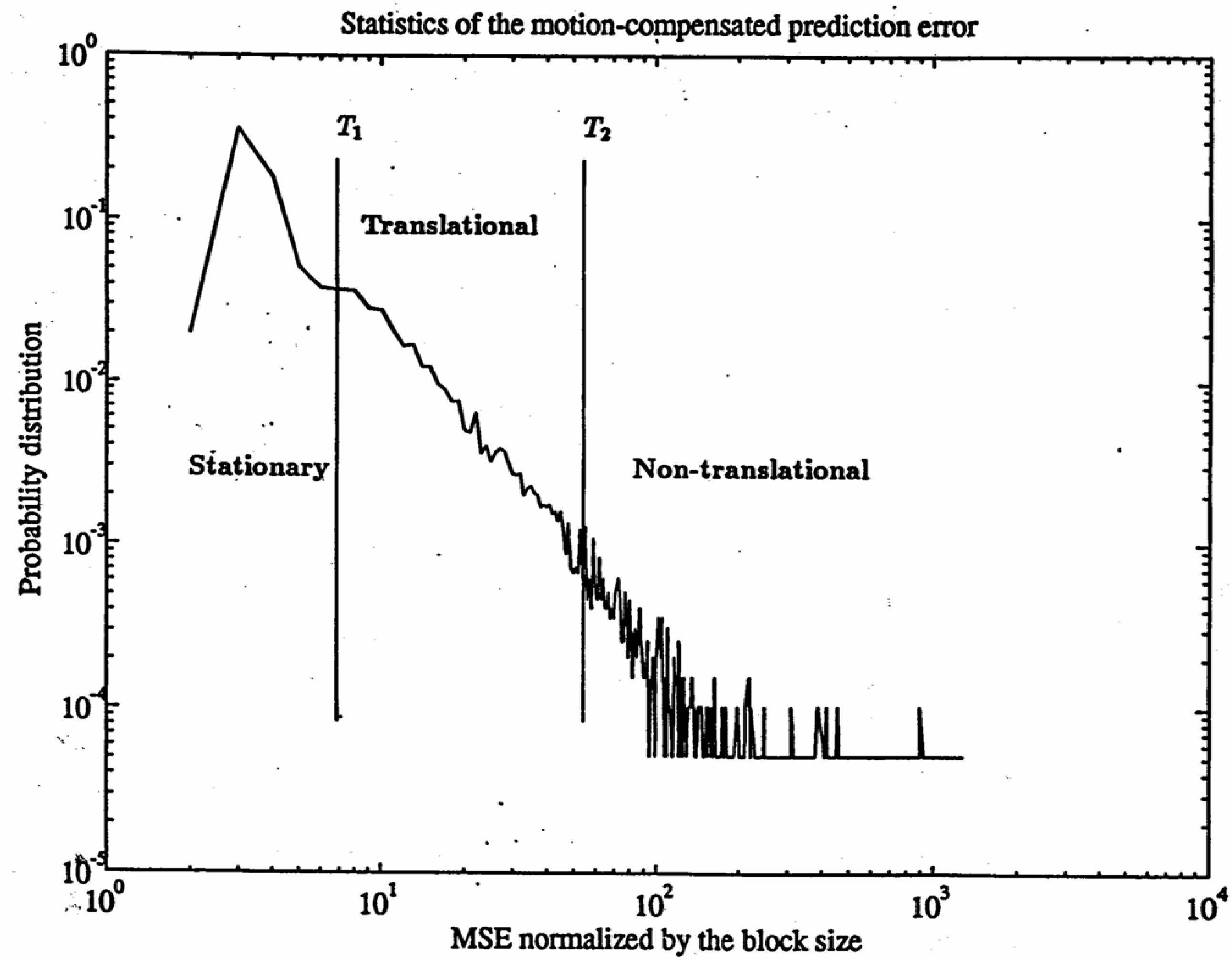
Hilbert scan

LRU Replenishment



A: Raster: Fixed
B: Raster: Variable
C: Hilbert: Fixed
D: Hilbert: Variable

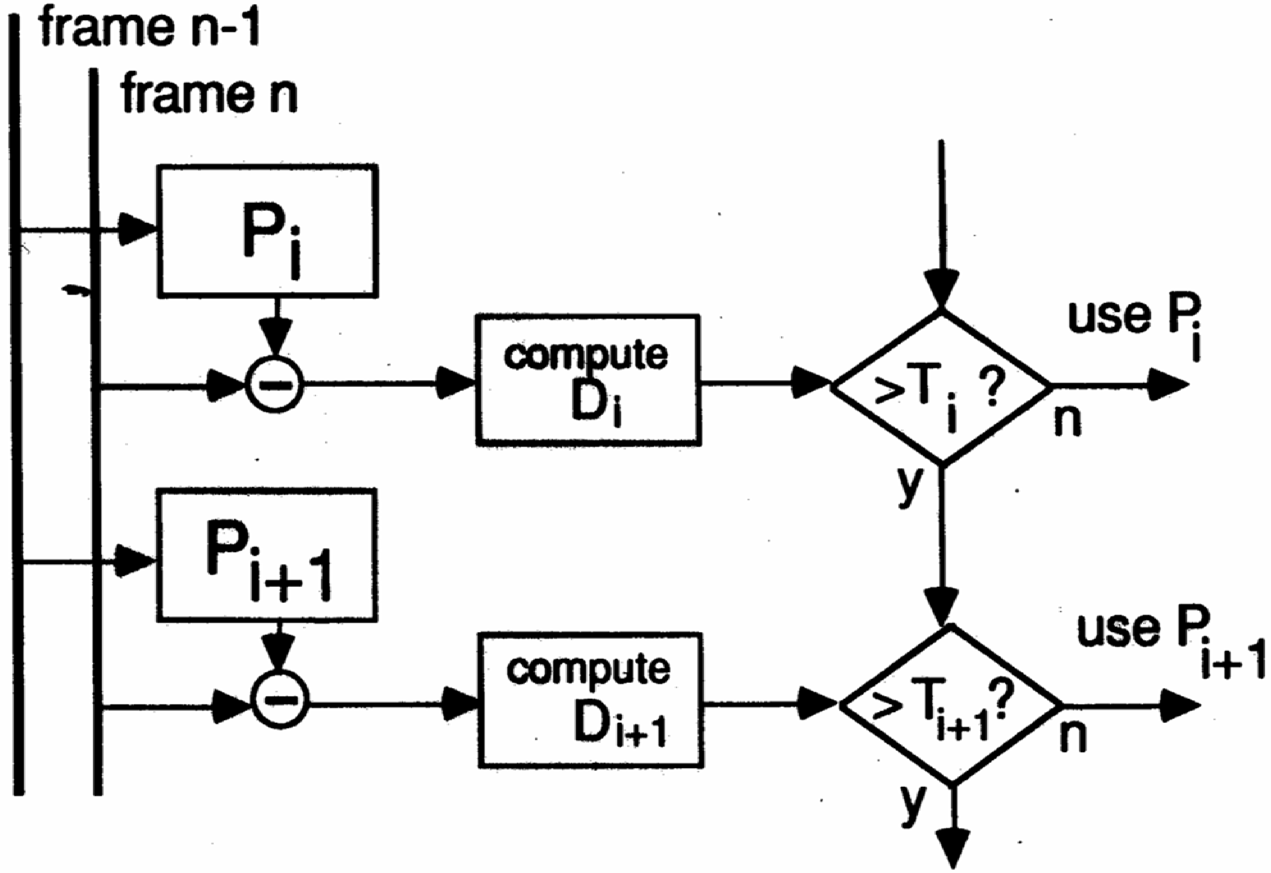
Hierarchical Video Encoding



Georgia Institute of Technology

CSIP Center for Signal & Image Processing

Hierarchical Encoding



10 Layer Video Encoder

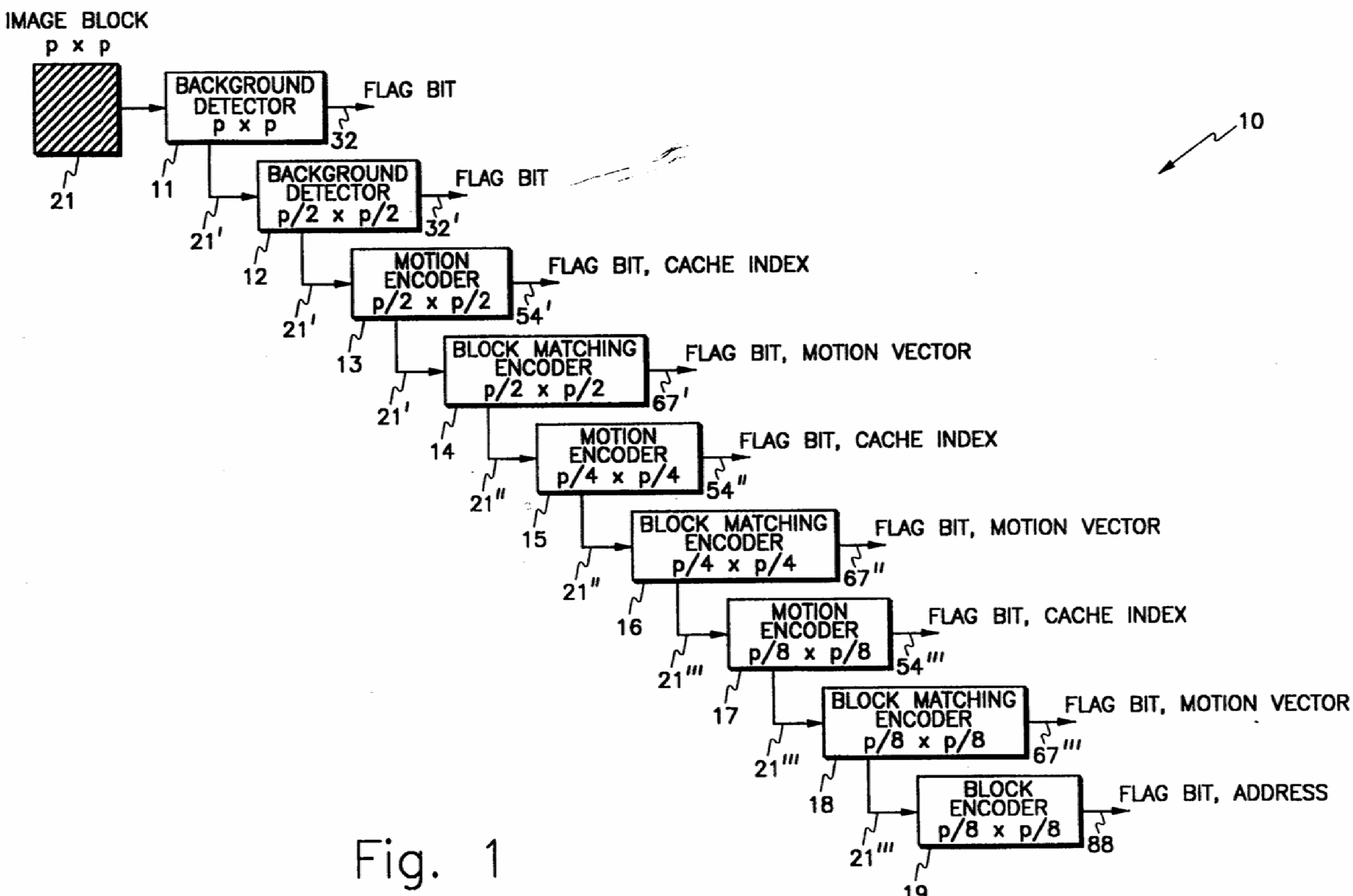
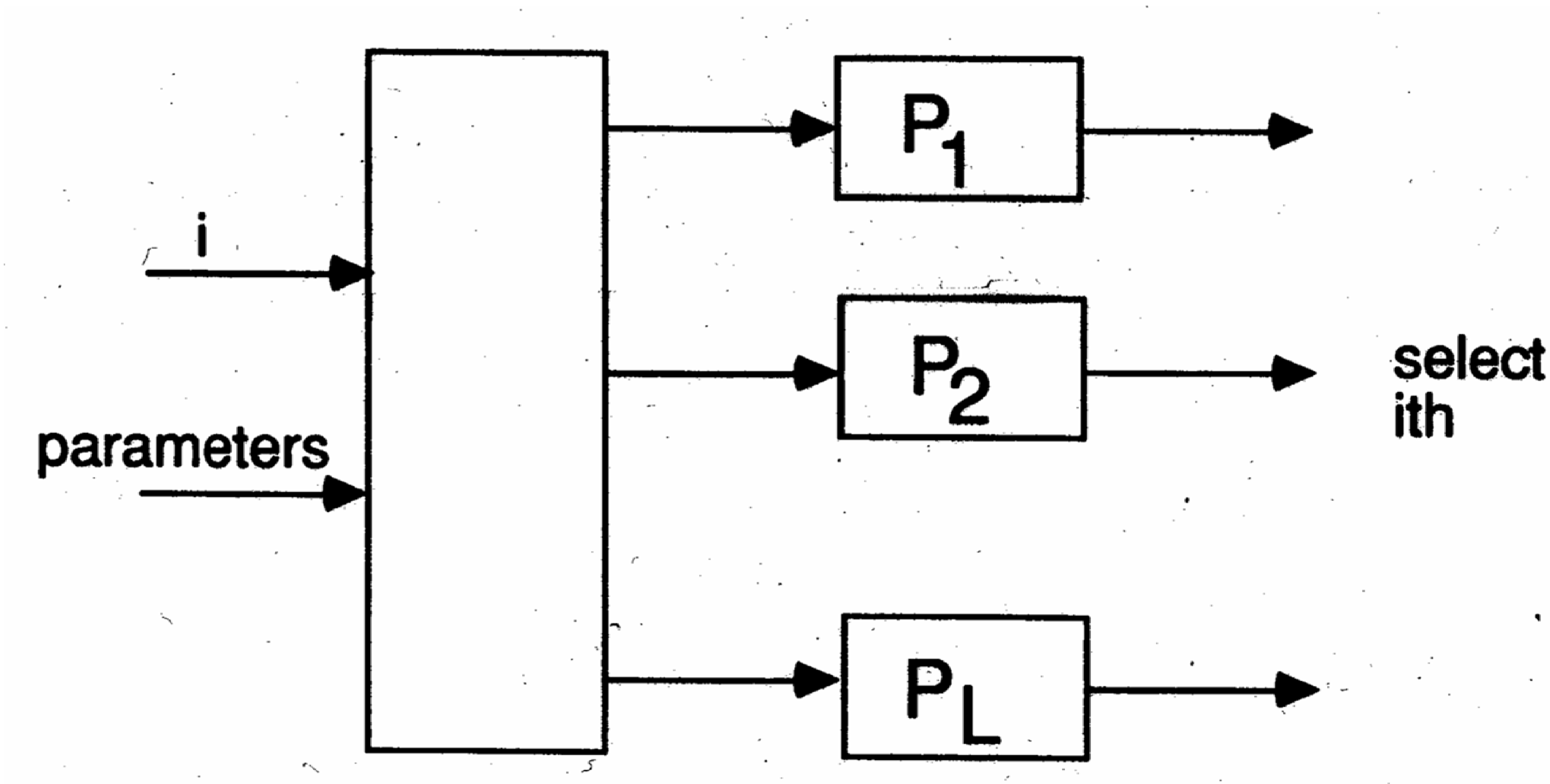


Fig. 1

Georgia Institute of Technology

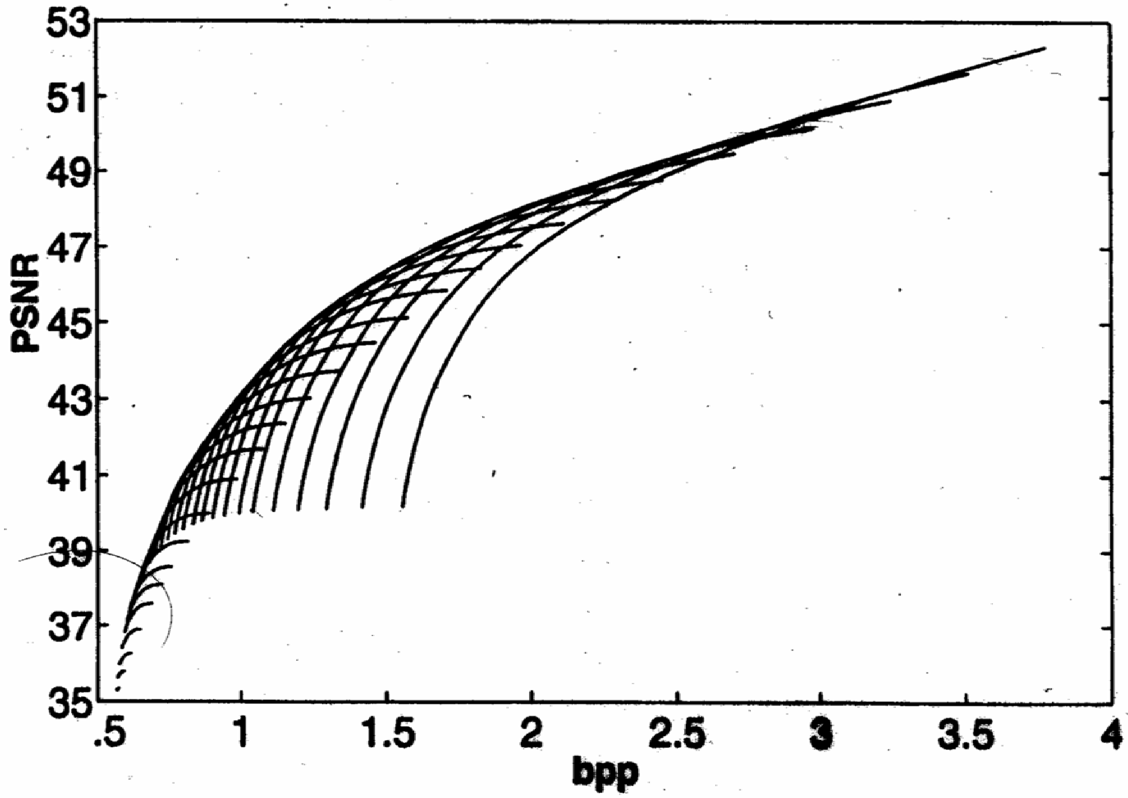
CSIP Center for Signal & Image Processing

Decoder



Threshold Selection

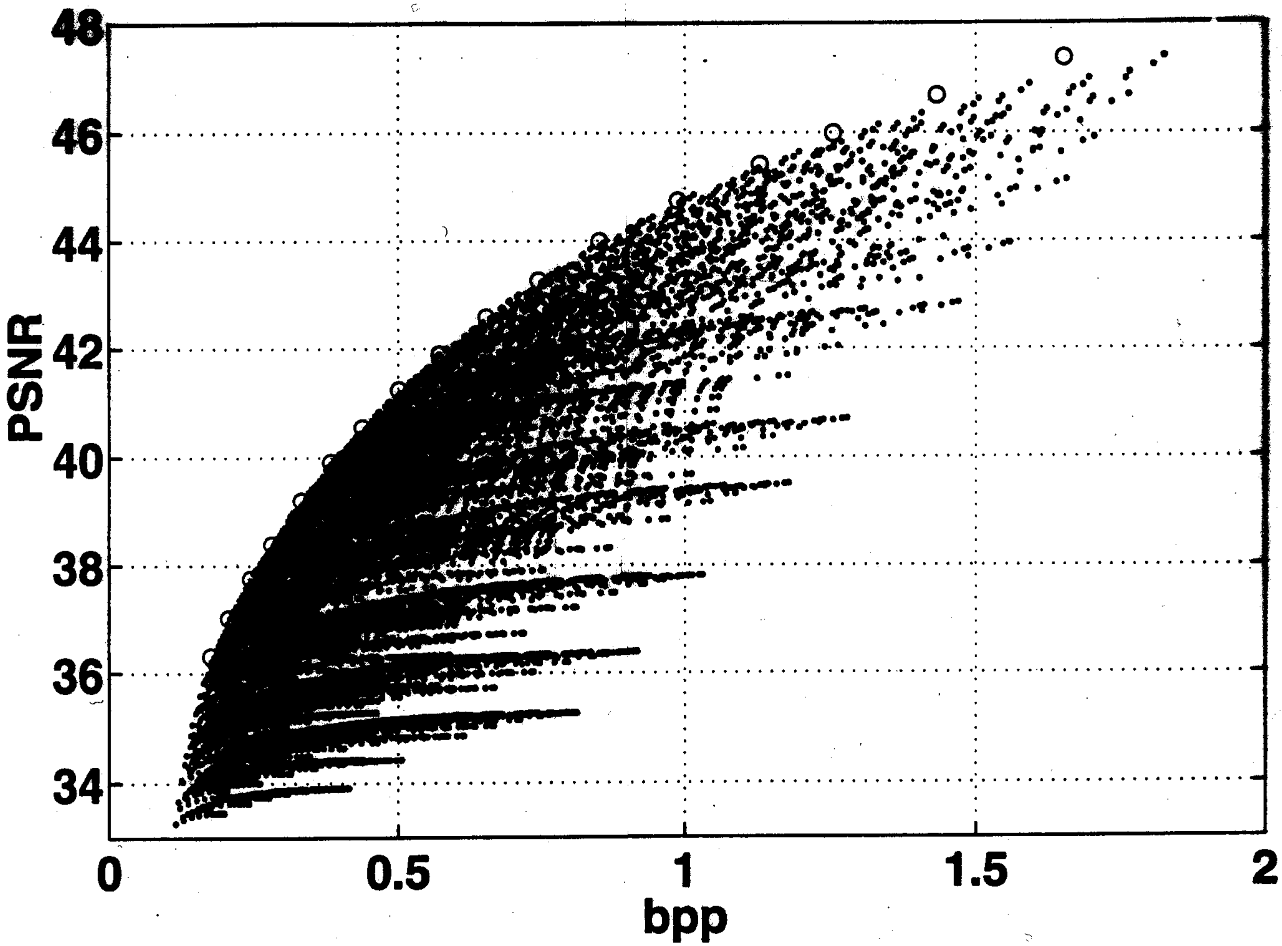
Setting the thresholds to be equal does *not* yield optimal rate-distortion performance.



Georgia Institute of Technology

CSIP
Center for Signal & Image Processing

Threshold Selection--Video



Georgia Institute of Technology

CSIP Center for Signal & Image Processing

Performance

- QCIF 10 frames/sec; 16 kbps

	Min Y PSNR	Avg Y PSNR
MissA (color)	34.38	35.53
MissA (mono)	36.02	36.64
Claire (color)	33.72	34.41
Claire (mono)	33.98	35.42
Sales (color)	28.41	30.18
Sales (mono)	29.34	31.00

Dynamic Codebook Reordering

- Extension of Cache VQ
- Codewords arranged by decreasing likelihood given the past selection(s).
  - ➔PMF of the address selected is highly skewed
- Selected addresses are entropy coded.

Principle

- At each t, DCR reorders the codebook vectors in increasing order of a dissimilarity measure between $Q[x(t)]$ and all other vectors in C.
- Re-ordering can be duplicated at the decoder without side information.

Example

- Let K=4; Codebook $\{C_0, C_1, C_2, C_3\}$
- At t=0 C_2 is selected, and 2 is transmitted.
- Let $D(C_2, C_2) < D(C_2, C_3) < D(C_2, C_1) < D(C_2, C_0)$
- At t=1, codebook is $\{C_2, C_3, C_1, C_0\}$
- The next symbol to be used is likely to be close to 0.

Dynamic Index Map

- Reordering the codebook can be expensive
- Define a dynamic index map $\Psi(k,t)$
- $\Psi(k,t)$ = position of C_k at time t.
- At time t, instead of sending k, we send $\Psi(k,t)$.

VQ Encoder with DCR

1. Codebook Search: Determine the "best match" codebook vector.

   $Q[x(t)] = C_k$

2. Dynamic Index Map: Find the reordered index

   $\Psi(k,t)$

3. Dynamic Codebook Reordering: determine $\Psi(k,t+1)$

Dynamic Codebook Reordering

1. Calculate

$$\delta(k,t) = D(Q[\mathbf{x}(t)], C_k) \text{ for } k=0,1,\ldots,K\text{-}1$$

2. Arrange δ in ascending order

$$\delta(k_0,t) \leq \delta(k_1,t) \leq \ldots \leq \delta(k_K,t)$$

3. $\Psi(j,t+1) = k_j$ for $j = 0,1,2,\ldots,K\text{-}1$

VQ Decoder with DCR

1. Inverse Dynamic Index Map

$$k = \Psi^{-1}(i(t), t)$$

2. Reconstruct using C_k

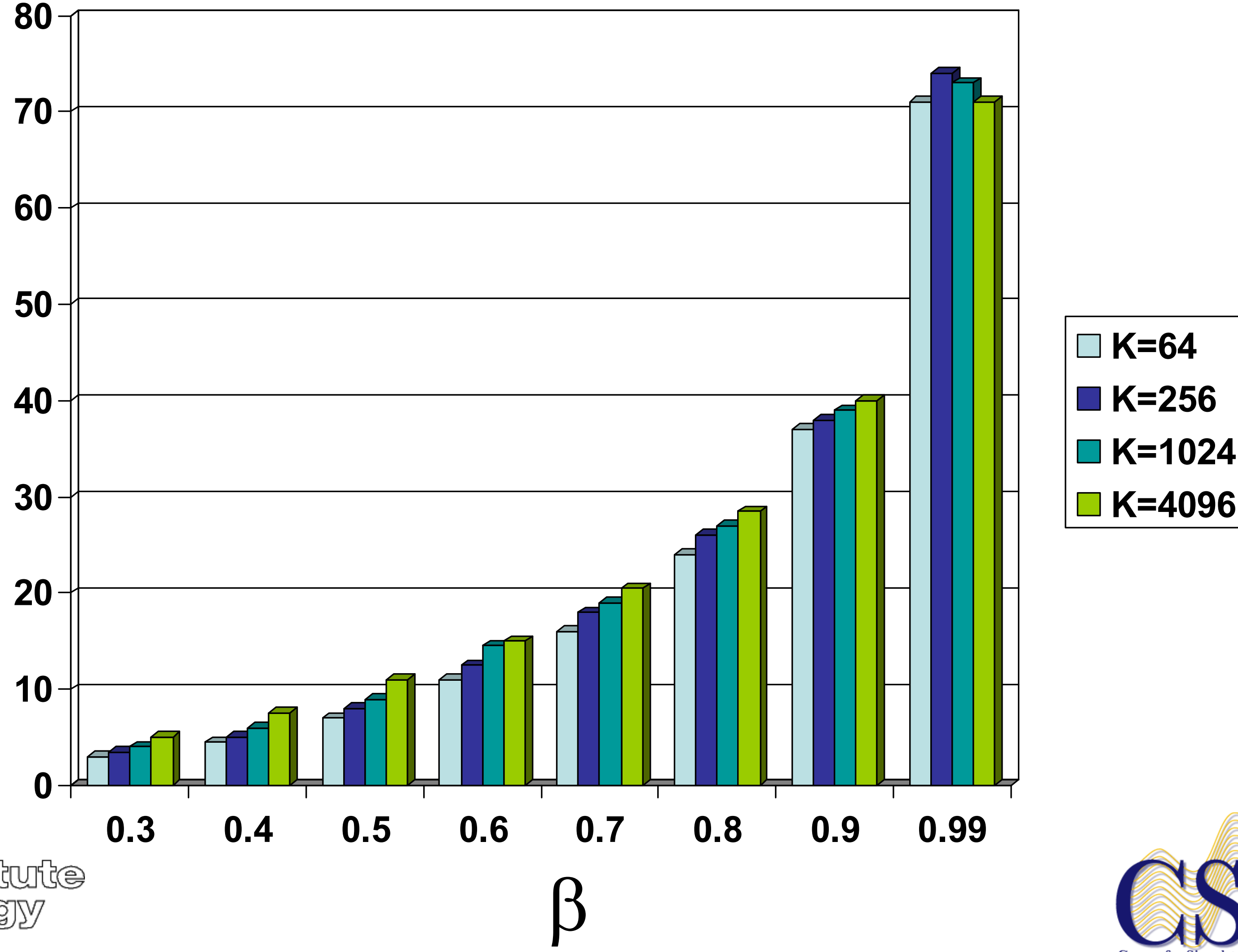
3. Update Dynamic Index Map

$$\Psi(i(t), t+1)$$

Lossless and Lossy Versions

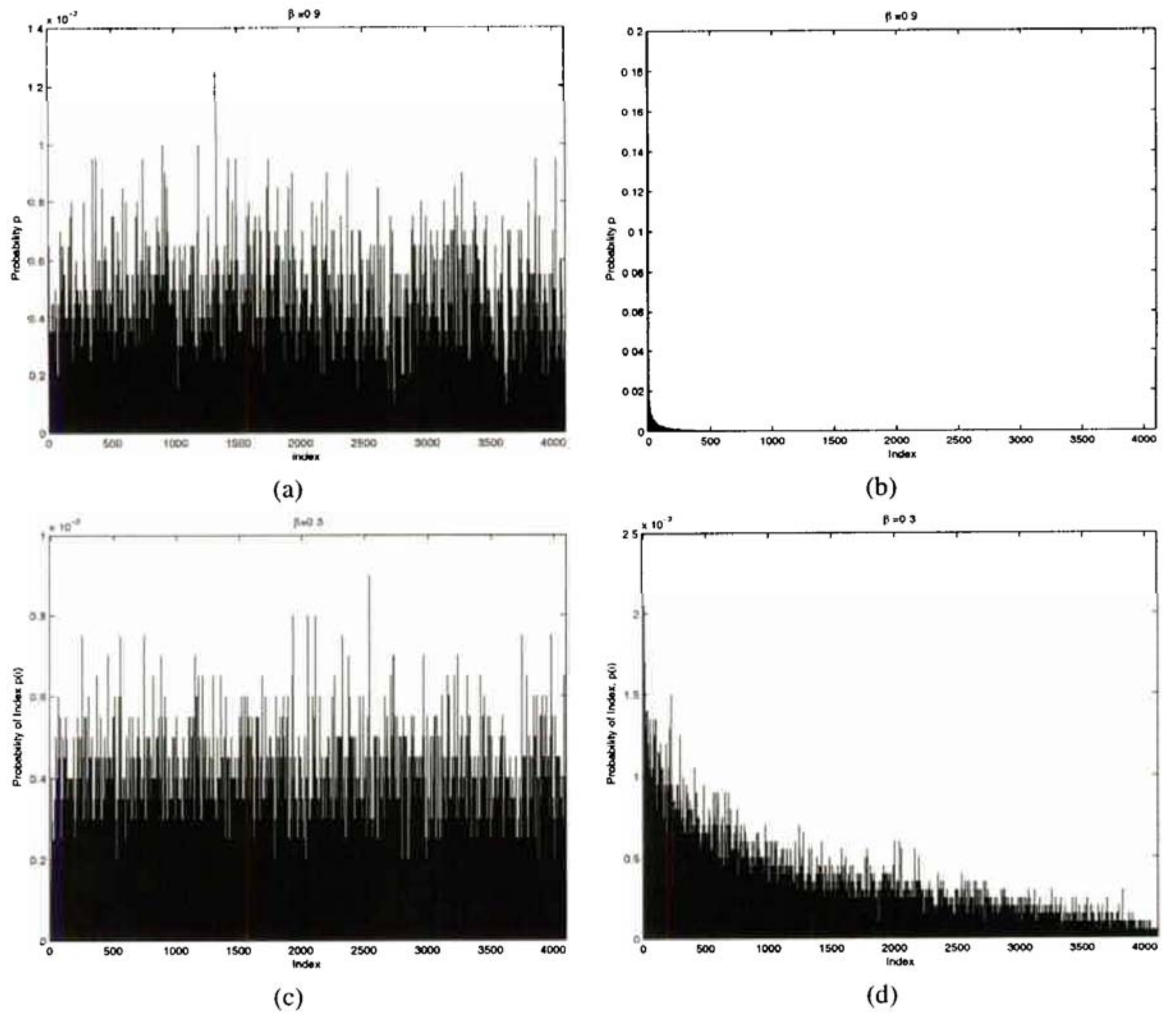
- In the lossless mode the entire codebook is searched
  - Entropy is reduced
  - Distortion is unchanged
- In the lossy mode a threshold defines a "good enough" fit
  - Entropy is reduced
  - Search time is reduced
  - Average distortion increased slightly

% Reduction in Entropy



Empirical pmfs

β=0.9



β=0.9
DCR

β=0.3

β=0.3
DCR

Georgia Institute of Technology

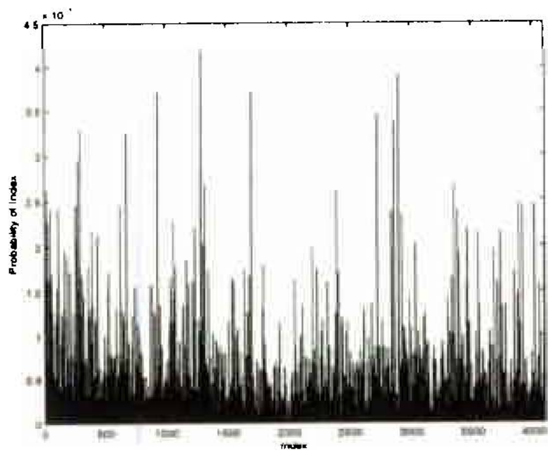
CSIP Center for Signal & Image Processing

MELP Speech Coder

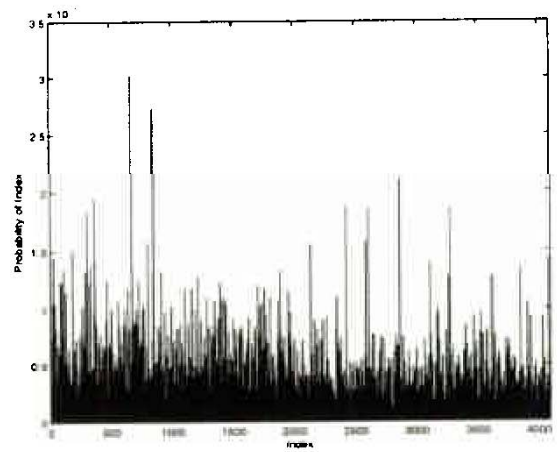
- DoD standard for low bitrate speech
  - 10-30 ms frames
  - Autoregressive system model coded as line spectral pairs
  - Excitation signal broken into subbands
  - Each subband contains a mixture of periodic and noise-like components

PMFs of MSVQ for LSFs

Stage 1
No DCR



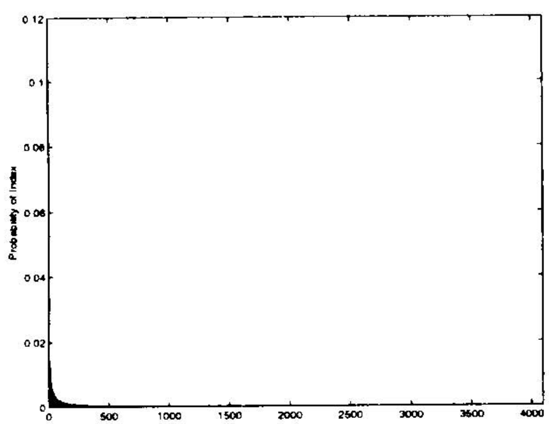
(a)



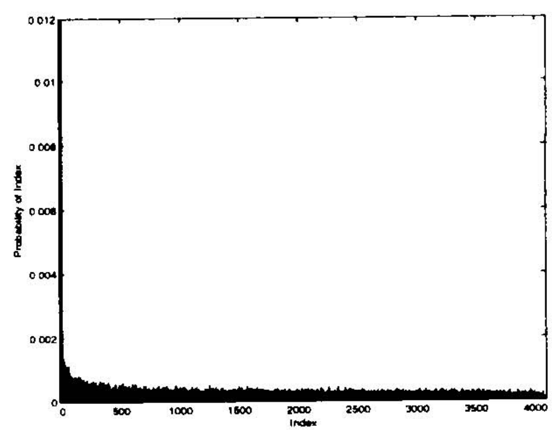
(b)

Stage 2
No DCR

Stage 1
DCR



(c)



(d)

Stage 2
DCR

Bit Allocation

Parameter	DOD Standard MELP		MELP with DCR	
	voiced	unvoiced	voiced	unvoiced
LSFs	25	25	16.63	16.63
Gain	8	8	6.51	6.51
Pitch	7	7	3.67	0
Bandpass Voicing	5	0	2.60	0
Fourier Magnitudes	8	0	6.06	0
Aperiodic Flag	1	0	1	0
Error Protection	0	13	0	0
Total	54	54	36.47	23.14

Dynamic Codebook Reordering Gain

- DoD standard MELP: **2400 bps** fixed
- MELP with DCR: **1500 bps** variable
- Identical output samples !