Context-Based Multiple Description Wavelet Image Coding

Dror Porat

Department of Electrical Engineering
Technion – Israel Institute of Technology

M.Sc. Thesis

Supervised by: Prof. David Malah

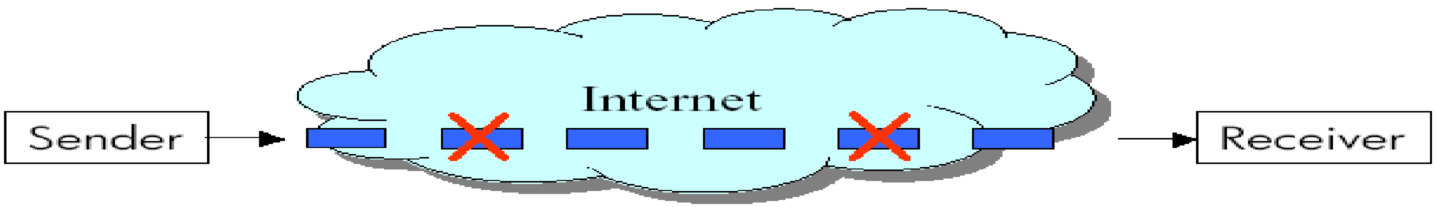
September, 2009

Outline

- Fundamentals of Multiple Description (MD) coding
- Framework: MD coding via polyphase transform and selective quantization
- Proposed system:
  Context-based MD wavelet image coder
  - Motivation
  - Detailed description
- Experimental results
- Summary and future directions

Fundamentals of MD Coding: Introduction

- Where does an SD (Single Description) coding system go wrong?



  - Packet losses!
    - Intolerable retransmission delay
    - No feedback channel
    - Order must be maintained (layered coding)

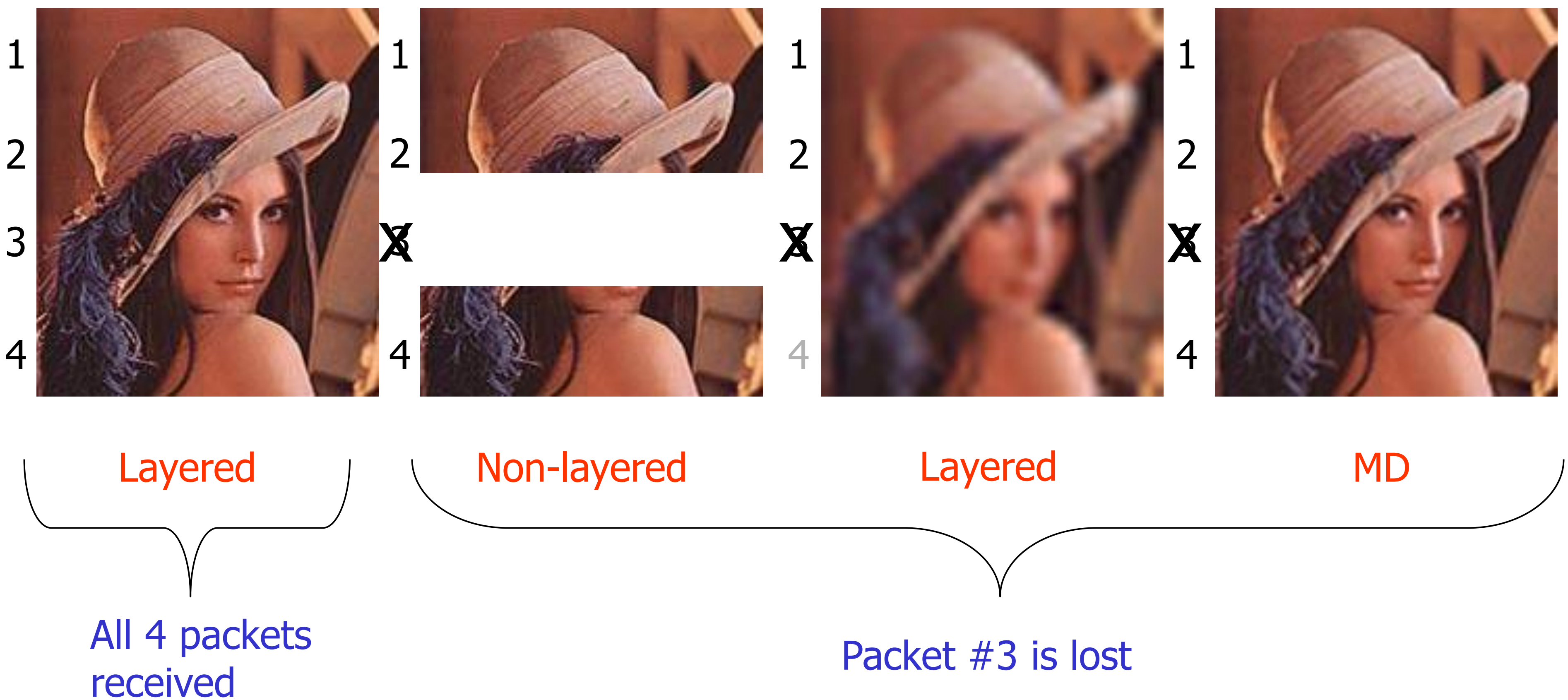
Fundamentals of MD Coding: Introduction (cont.)

- Purpose: Provide error resilience to information transmitted on lossy networks (e.g., the Internet)
- Possible solution:
  MD coding
  - Represent the information source with several descriptions
  - The source can be approximated from any (non-empty) subset of the descriptions

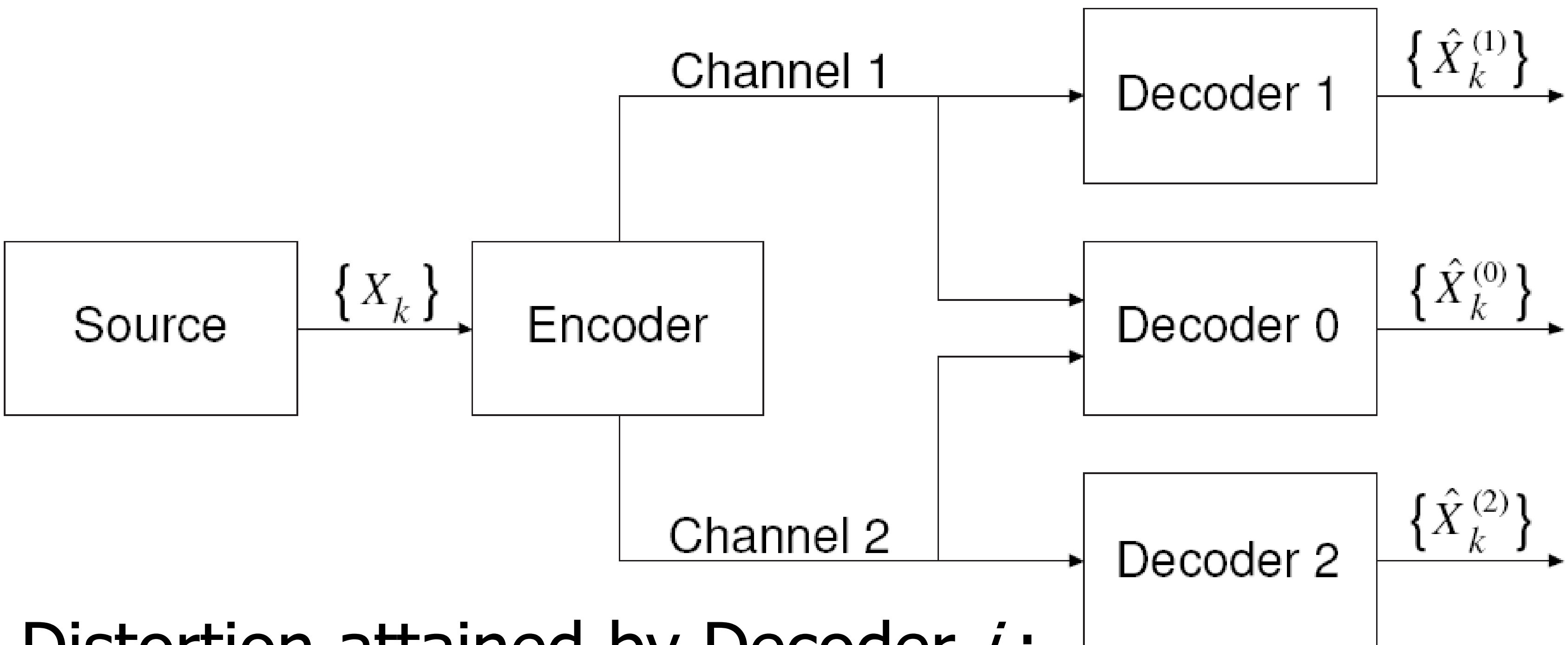
⇒ Makes all received descriptions useful

- Both encoder and decoder are involved (different from post processing-based error concealment)
- Caveat: No free lunch! (Redundant representation)

Fundamentals of MD Coding: Example – MD Coding vs. Layered Coding



Fundamentals of MD Coding: Scenario for MD Coding (Two Descriptions)



- Distortion attained by Decoder *i* :

$$D_i = \frac{1}{N}\sum_{k=1}^{N} E\left[d\left(X_k, \hat{X}_k^{(i)}\right)\right], \; i = 0, 1, 2$$

N – Number of source symbols

Fundamentals of MD Coding: Information Theoretic Aspects

- MD rate distortion region (MD region):
  Closure of the set of achievable quintuples (R_1, R_2, D_0, D_1, D_2)
- Achievable quintuples (not all) by [El Gamal and Cover, 1982]
- The MD region for a memoryless Gaussian source with variance σ^2 and squared error distortion measure:

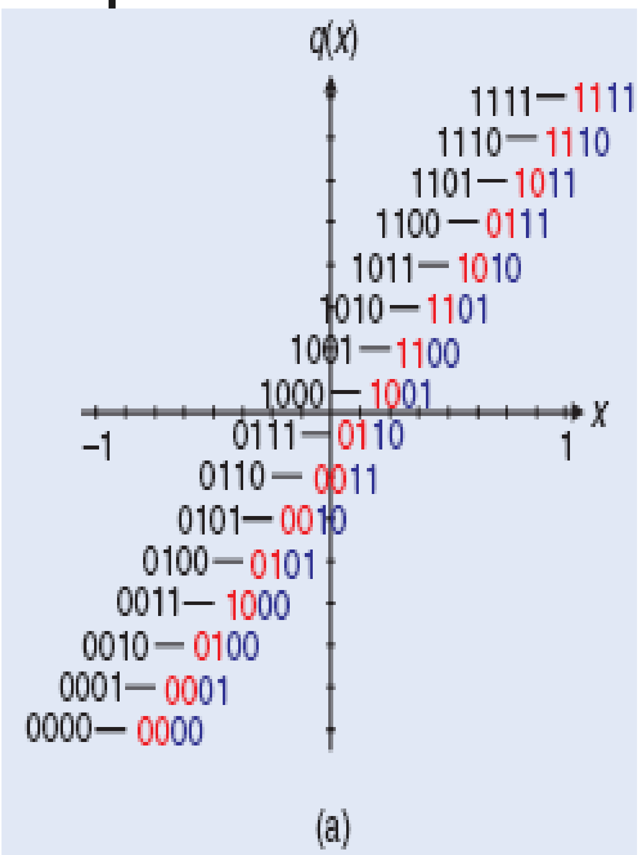
$$D_i \geq \sigma^2 2^{-2R_i}, \quad i = 1, 2$$

$$D_0 \geq \sigma^2 2^{-2(R_1+R_2)} \cdot \gamma_D$$

where $\gamma_D = \begin{cases} 1 & \text{, if } D_1 + D_2 > \sigma^2 + D_0 \\ \frac{1}{1-\left(\sqrt{(1-D_1)(1-D_2)}-\sqrt{D_1 D_2 - 2^{-2(R_1+R_2)}}\right)^2} & \text{, else} \end{cases}$

Fundamentals of MD Coding: MD Scalar Quantization (MDSQ)

- Example: Communicate a single real number $x \in [-1, 1]$



4-bit quantizer (From [Goyal, 2001])

Two 3-bit quantizers (Total rate: 6 bits)

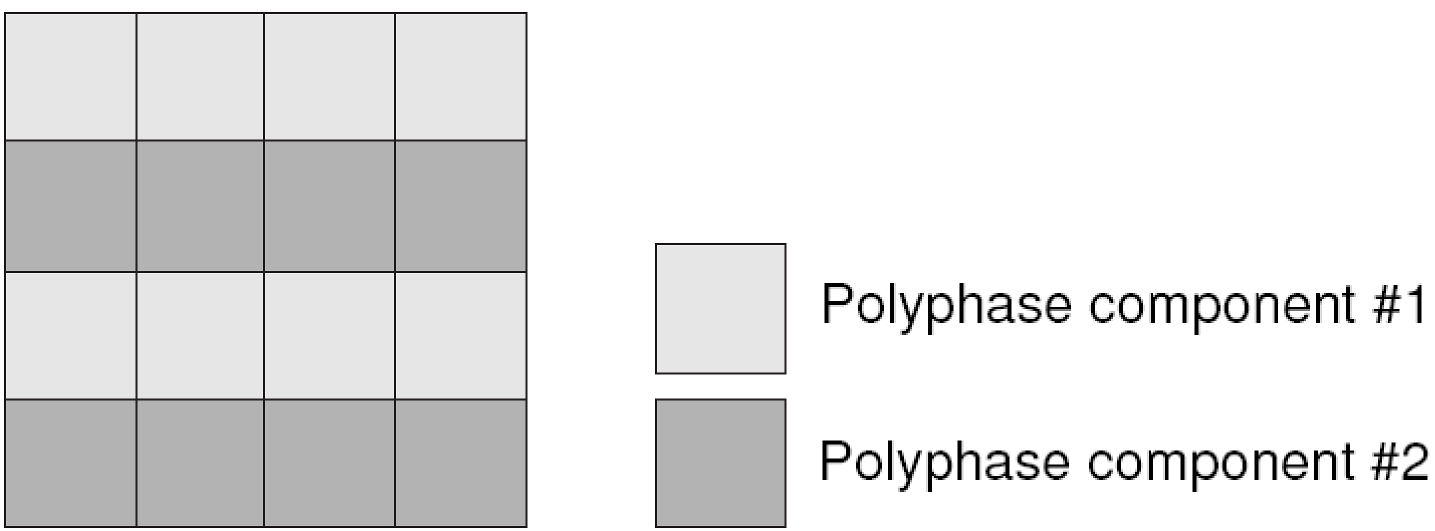
"Complicated" MDSQ (Total rate: ~5.2 bits)

Framework: MD Coding via Polyphase Transform and Selective Quantization

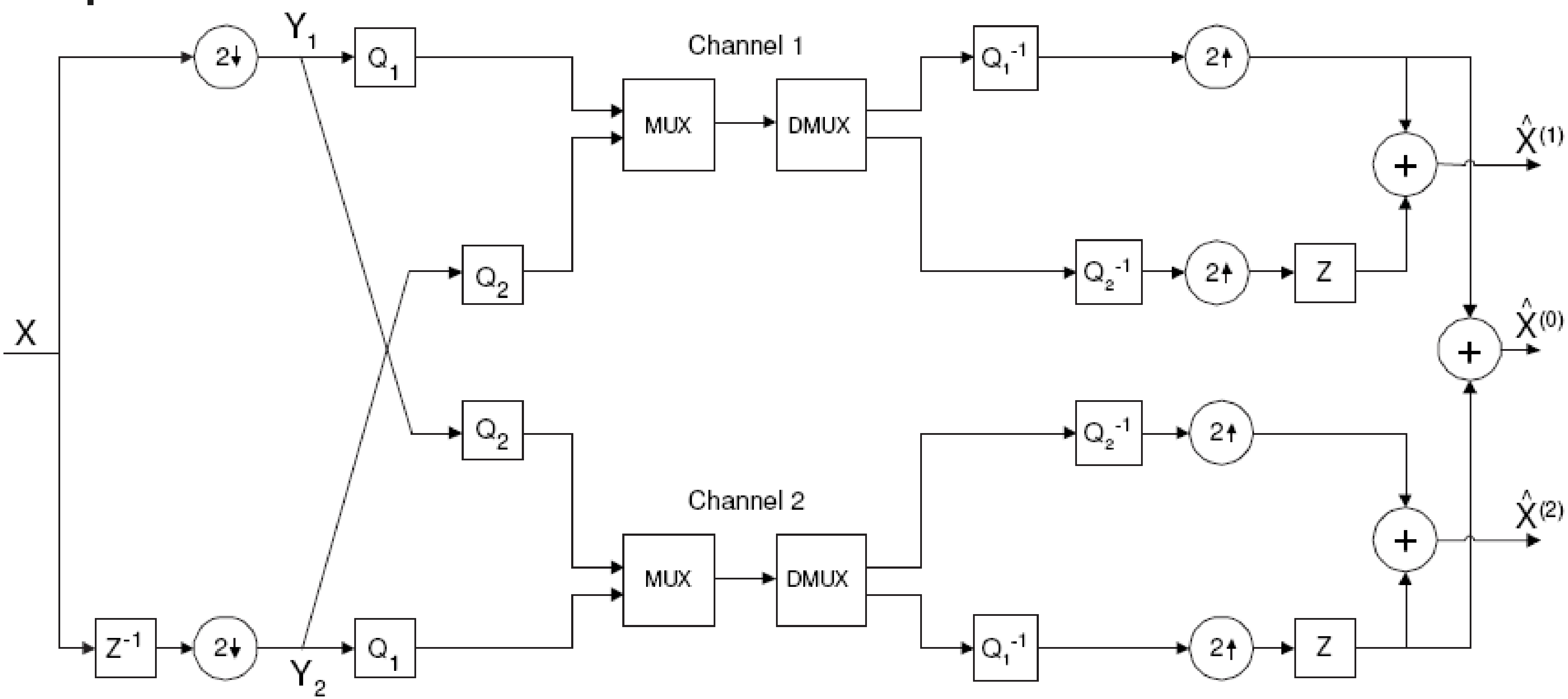
- Proposed by [Jiang and Ortega, 1999]
- Explicitly separates description generation and redundancy addition
  - ⇒ Reduced complexity of design and implementation
- Enables simple generation of descriptions of equal rate and importance (balanced case)
  - ⇒ Well suited to communication systems with no priority mechanisms (e.g., the Internet)

Polyphase Transform-Based MD Coding: The Polyphase Transform

- Polyphase transform:
  Decomposition to polyphase-like components
- Example – plain polyphase transform:



Polyphase Transform-Based MD Coding: System Outline

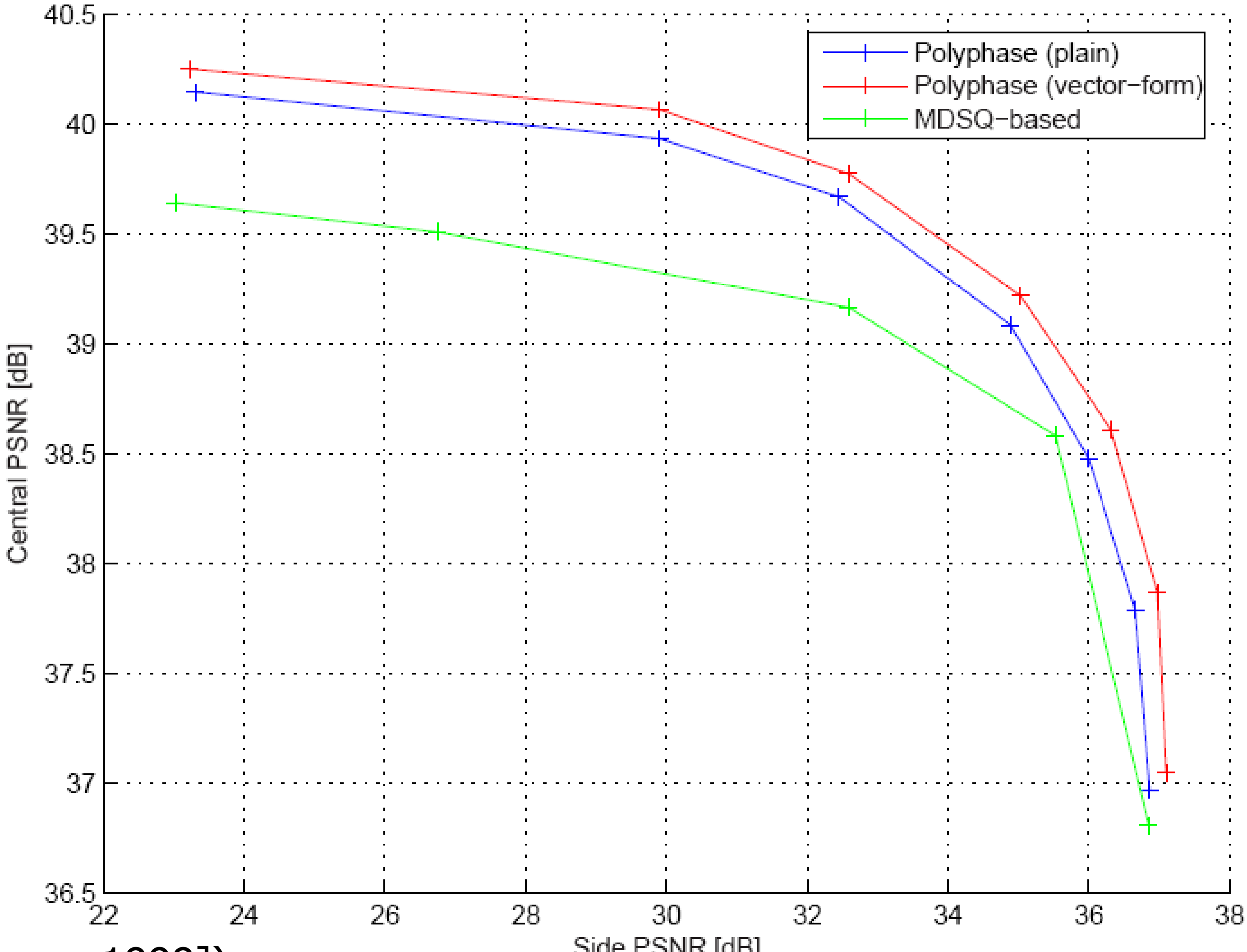


- For correlated input data (e.g., an image), a preliminary decorrelating transform is required

Polyphase Transform-Based MD Coding: Experimental Results

- Wavelet transform used for decorrelation
- Polyphase transform – two alternatives:
  - Plain
  - Vector-form: Groups coefficients in different subbands corresponding to the same spatial location (similar to the zerotree structure)
- SPIHT used for quantization and entropy coding

Polyphase Transform-Based MD Coding: Experimental Results (cont.)



Lena (grayscale)

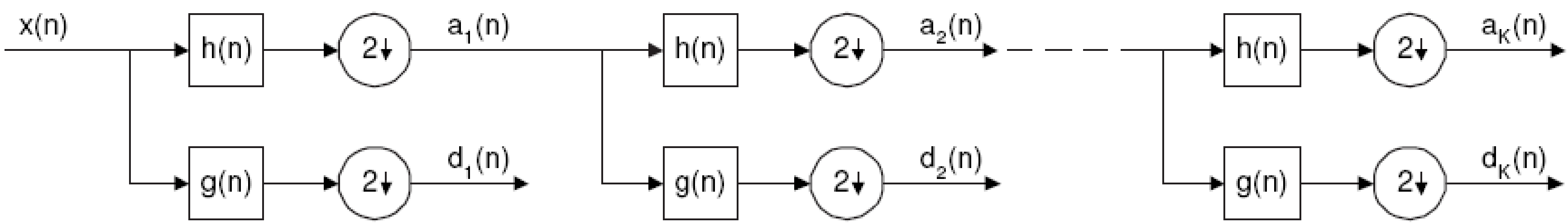
512x512 pixels

Total rate = 1 bpp

(From [Jiang and Ortega, 1999])

Context-Based MD Wavelet Image Coding: Wavelet Background

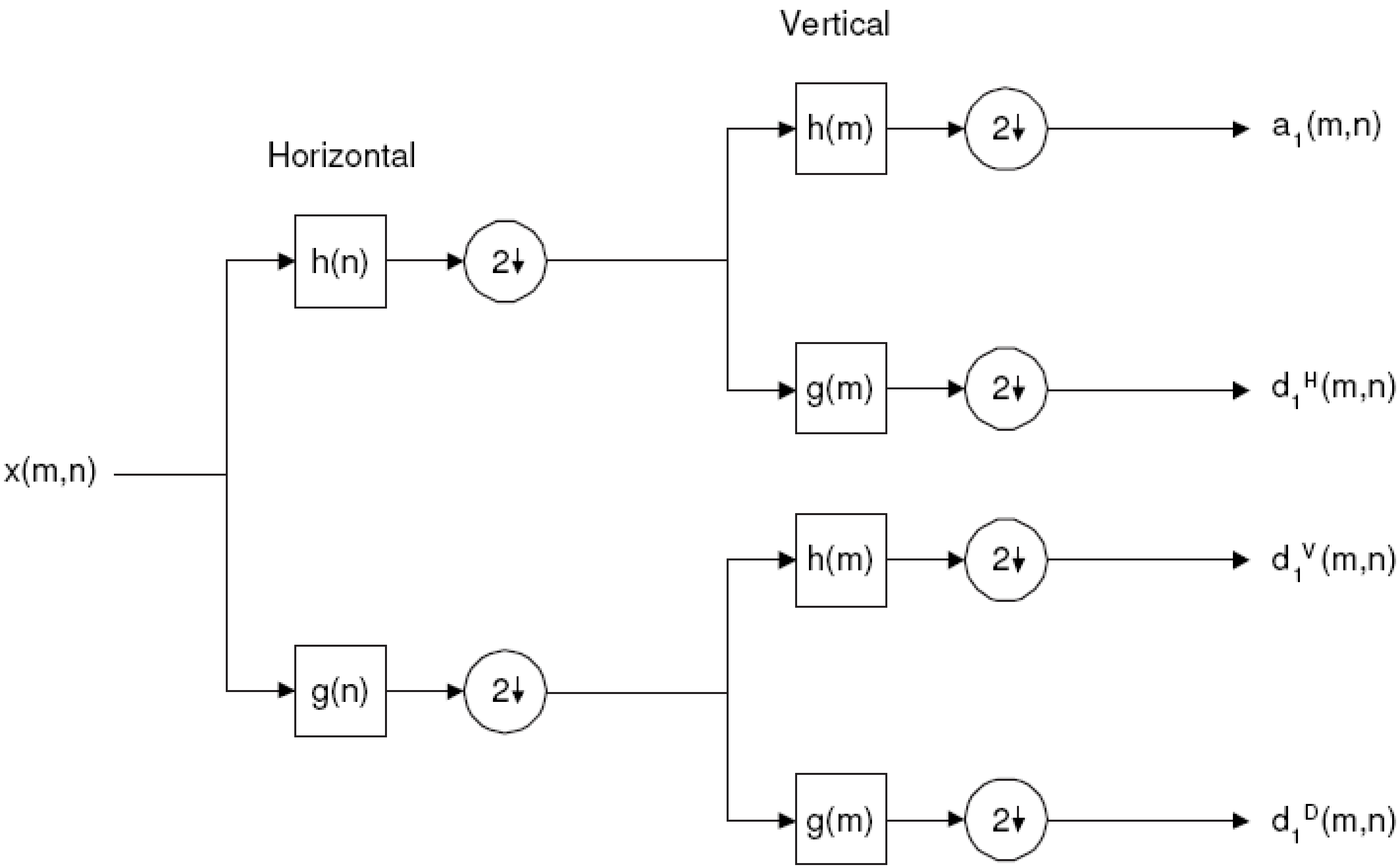
- K-level 1-D discrete wavelet transform (DWT):

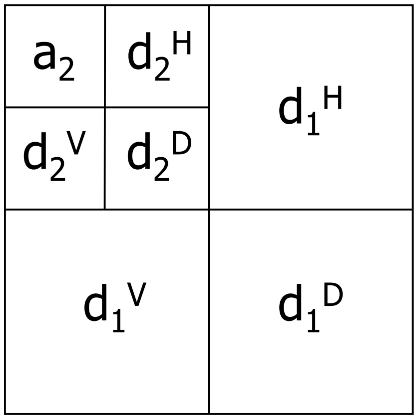


- Approximation coefficients: $\{a_K(n)\}$
- Detail coefficients: $\{d_i(n)\}$, $i \in \{1,\ldots,K\}$

Context-Based MD Wavelet Image Coding: Wavelet Background (cont.)

- 2-D DWT using separable filters:



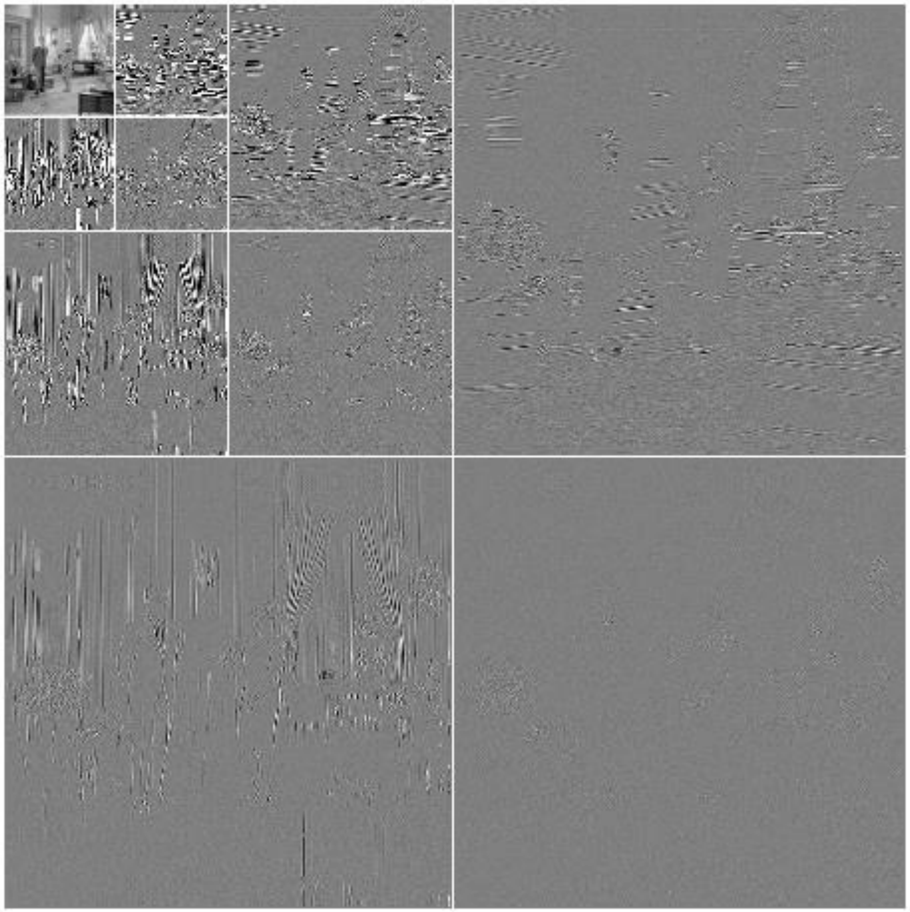


- Multiple levels (scales): Repeat on approximation

Context-Based MD Wavelet Image Coding: Wavelet Background (cont.)

- Example:

Original



Wavelet transform

Context-Based MD Wavelet Image Coding: Statistical Characterization of Wavelet Coeffs

- Wavelet transform provides good energy compaction (⇔ reducing the correlation amongst wavelet coeffs)
- First order statistics of detail wavelet coeffs successfully modeled using the Generalized Gaussian Distribution (GGD):

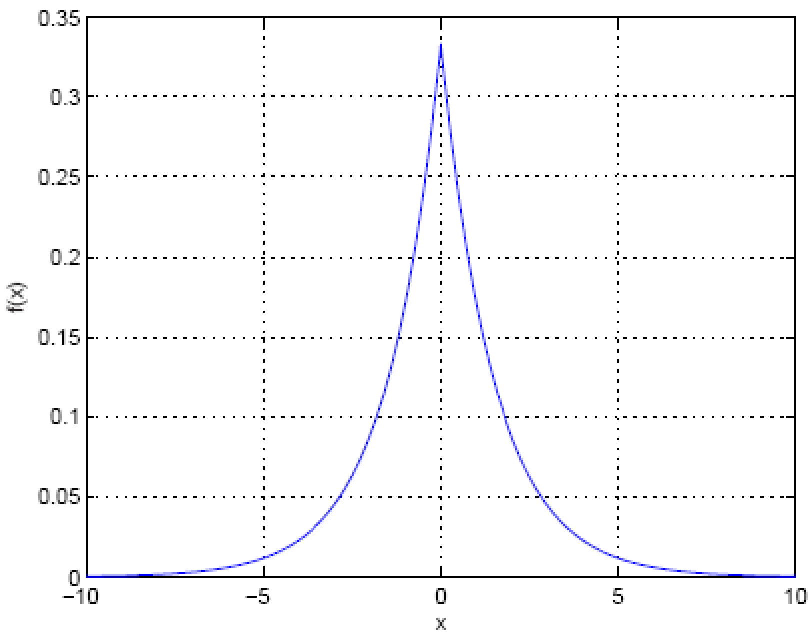
$$f_{s,r}(x) = \frac{1}{N(s,r)} e^{-|x/s|^r}$$

where $N(s,r) = 2s\Gamma(1/r)/r$ and $\Gamma(a) = \int_0^\infty t^{a-1} e^{-t}\, dt$

  - *r* = 2: Gaussian distribution
  - *r* = 1: Laplacian distribution (double exponential distribution)
  - $r \to \infty$: Uniform distribution (pointwise convergence on (-*s*,*s*))

Context-Based MD Wavelet Image Coding: Statistical Characterization of Wavelet Coeffs (cont.)

- Best fit of GGD for natural images: $r \in [0.5, 1]$ [Buccigrossi and Simoncelli, 1999]
  - ⇒ First order statistics of detail wavelet coeffs can be reasonably modeled using Laplacian distribution



$r = 1$, $s = 1.5$

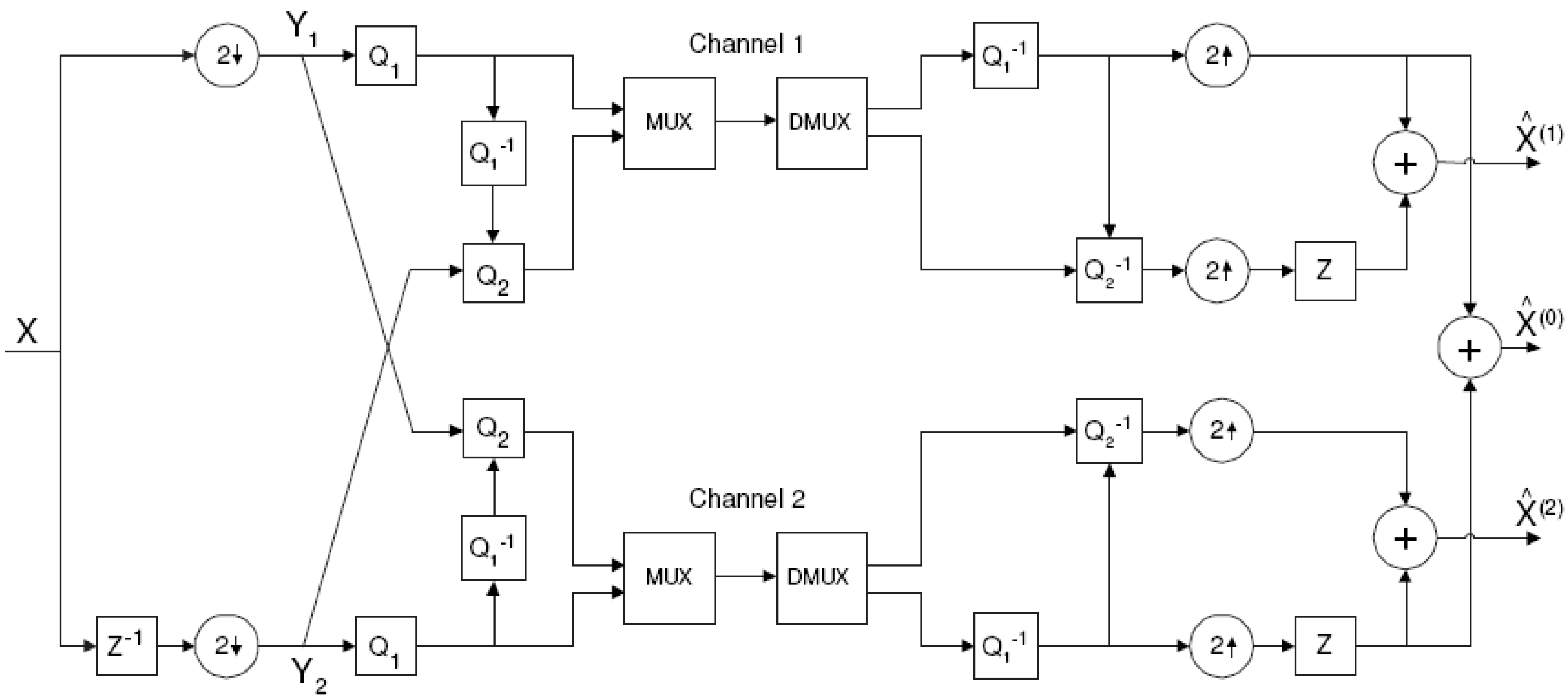
Context-Based MD Wavelet Image Coding: Statistical Characterization of Wavelet Coeffs (cont.)

Spatial and scale-to-scale dependencies:

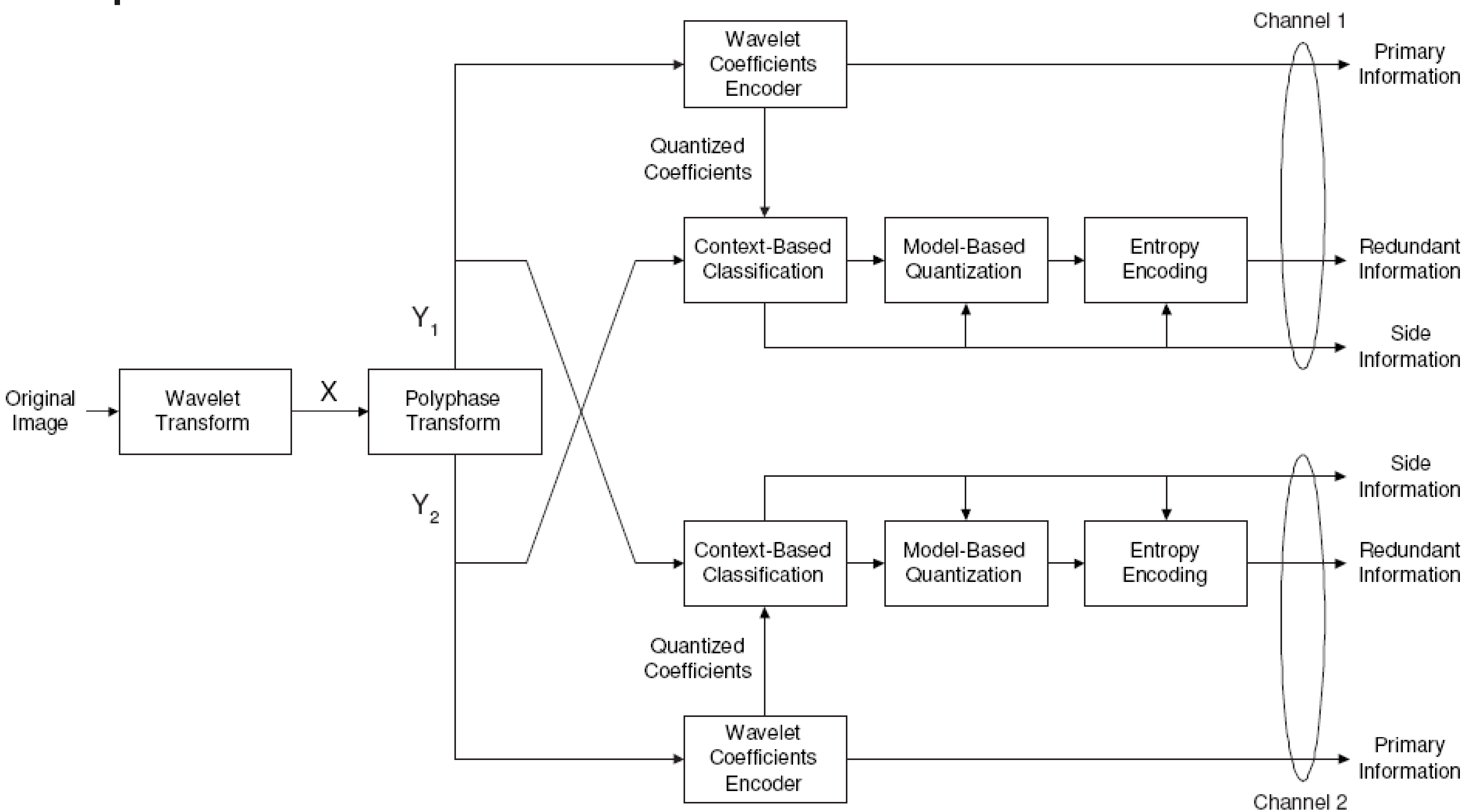
- Wavelet coefficients are not statistically independent (although approximately decorrelated)
- Potential conditioning neighbors:
- Dependencies are implicitly utilized by numerous image compression schemes (e.g., EZW, SPIHT)
- Linear predictor for the magnitude of a coefficient proposed by [Buccigrossi and Simoncelli, 1999]
  - Contribution to the mutual information between a coeff's magnitude and its predictor (in decreasing order): Local neighbors (“Left”, “Up”), parent, cousins, …

Context-Based MD Wavelet Image Coding: Proposed System Outline

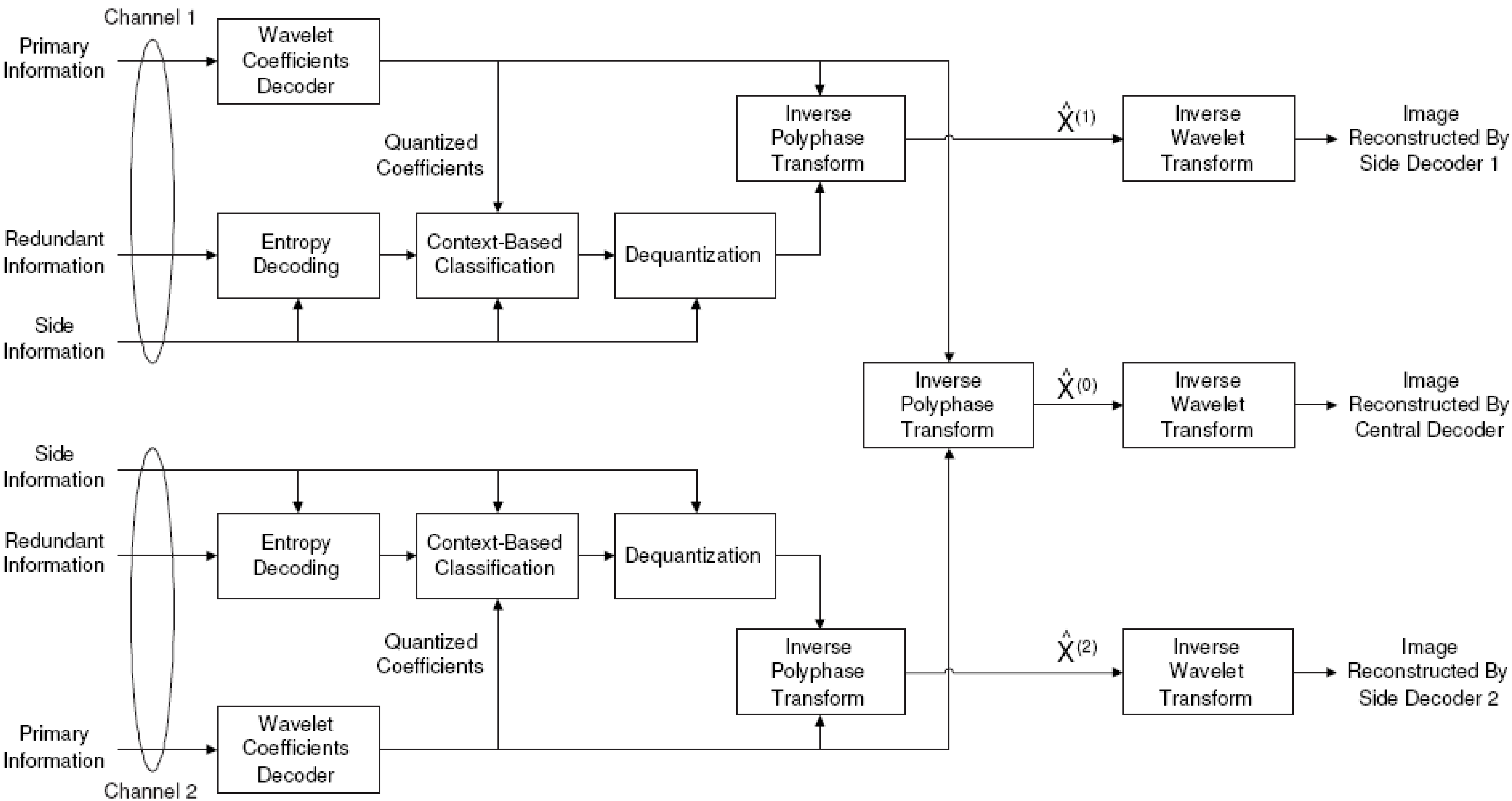
- Concept: Improve coding efficiency of Q_2 via utilization of contextual information from Q_1



Context-Based MD Wavelet Image Coding: Proposed MD Encoder – Block Diagram

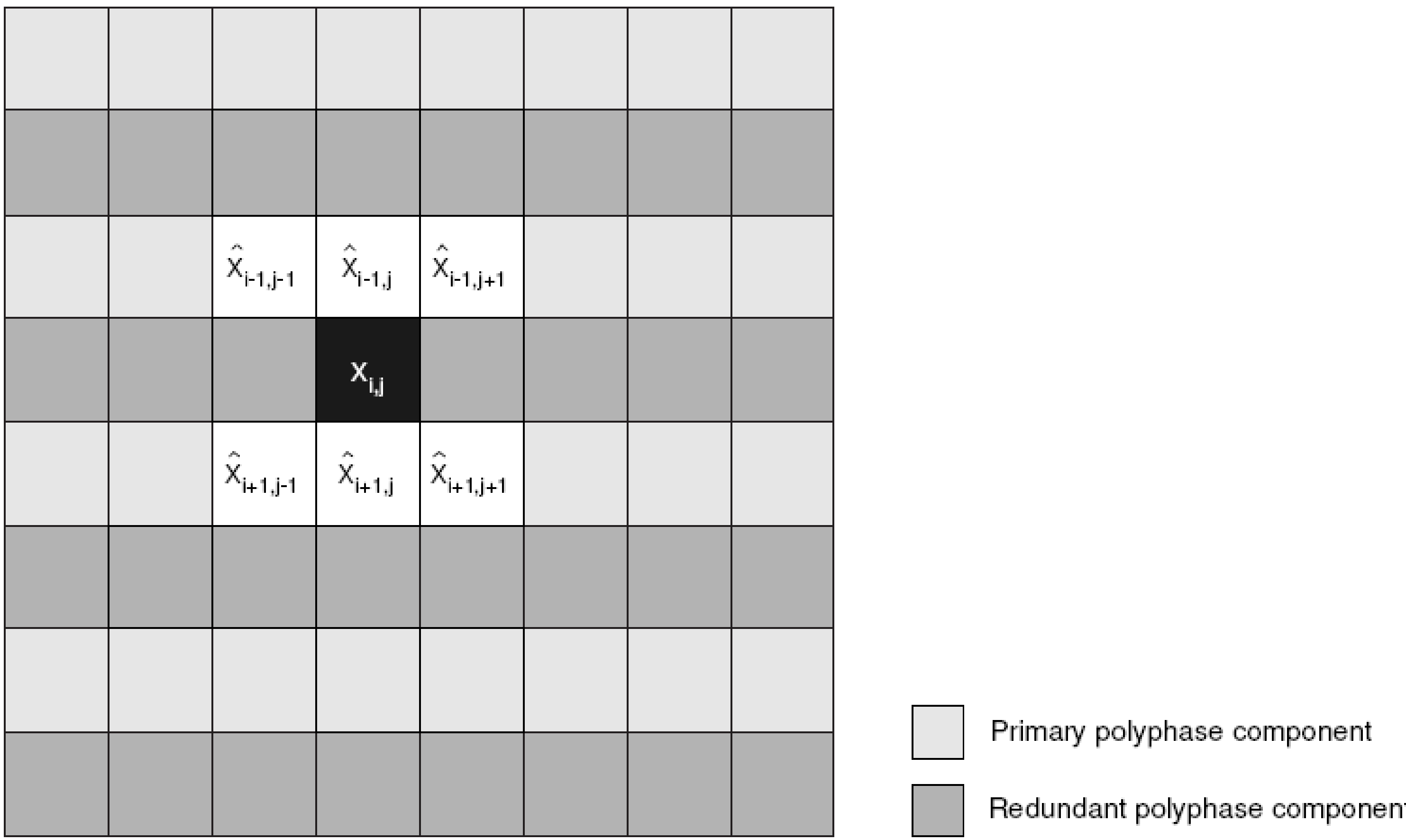


Context-Based MD Wavelet Image Coding: Proposed MD Decoder – Block Diagram



Context-Based MD Wavelet Image Coding: Context Formation

- Classification of the wavelet coefficient $X_{i,j}$ is based on the following context $C_{i,j}$ of quantized local neighbors:



Context-Based MD Wavelet Image Coding: Context-Based Classification

- Classification offers a potential increase in coding efficiency (quantization is adapted to the data)
- Penalty of forward classification is avoided (classification is based on quantized coefficients)
- Side information is transmitted for improved performance:
  - Classification thresholds (allowing to select a class for a coefficient given its context)
  - Source statistics of each class (each class is modeled using a parametric distribution)
- Avoids explicit characterization of statistical dependencies between neighboring wavelet coefficients
- Classification procedure inspired by [Yoo et al., 1999] (SD subband image coder)

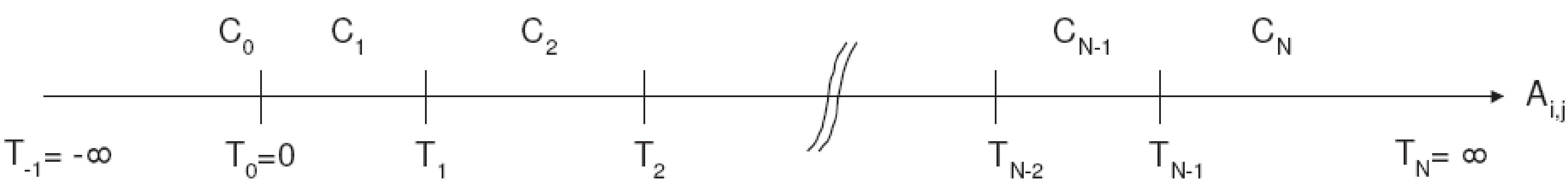
Context-Based MD Wavelet Image Coding: Classification Rule

- Purpose: Assign one of a finite number of classes to a coefficient $X_{i,j}$ given its context $C_{i,j}$
- Classification is based on a weighted average of the magnitudes of coefficients in $C_{i,j}$ ("Activity"):

$$A_{i,j} = a_1|\hat{X}_{i-1,j-1}| + a_2|\hat{X}_{i-1,j}| + a_3|\hat{X}_{i-1,j+1}| + a_4|\hat{X}_{i+1,j-1}| + a_5|\hat{X}_{i+1,j}| + a_6|\hat{X}_{i+1,j+1}|$$

where $\sum_k a_k = 1$

- E.g.: Weights are inversely proportional to the Euclidean distances of the corresponding coeffs in $C_{i,j}$ from $X_{i,j}$
- Classification rule (for set classification thresholds):



Context-Based MD Wavelet Image Coding: Classification Thresholds Design

- Purpose (for a given subband):
  - Determine the classification thresholds $T_1, T_2, \ldots, T_{N-1}$ from an initial set of N_0-1 candidate thresholds (where $N_0 > N$)
  - Goal: Maximization of the "classification gain" (coding gain due to classification, under certain simplifying assumptions)
- Model assumption: Coeffs in each class of each subband are drawn from a (zero-mean) Laplacian distribution

$$f_\lambda(x) = \frac{\lambda}{2} e^{-\lambda |x|}$$

  - Holds for approximation subband as well, due to employment of a DPCM-like prediction scheme (i.e., holds for prediction errors)
  - The Laplacian parameter for each class is estimated using MLE:

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} |x_i|} = \frac{1}{\frac{1}{n}\sum_{i=1}^{n} |x_i|}$$

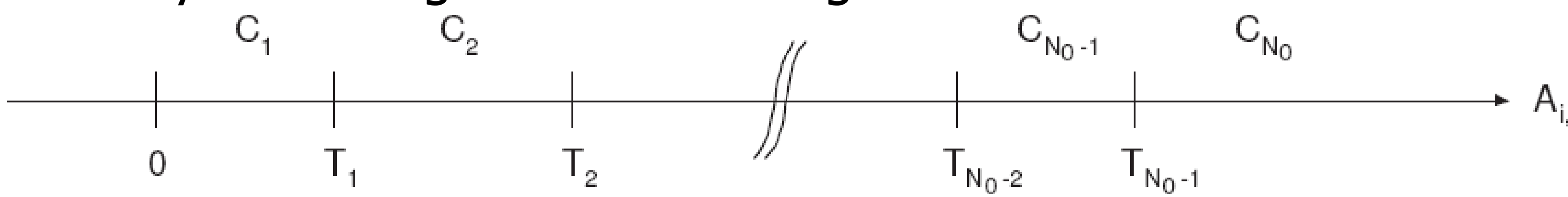
Context-Based MD Wavelet Image Coding: Classification Thresholds Design (cont.)

- Classification threshold design algorithm (given subband):

  Given: ▫ The desired total number of classes N+1

  ▫ N_0-1 strictly increasing initial thresholds (where N_0>N)

  - Classify all coeffs with zero activity to C_0 and estimate $\hat{\lambda}_0$ (MLE)
  - Classify remaining coeffs according to the classification rule



  and estimate (using MLE) the Laplacian parameters $\hat{\lambda}_1, \ldots, \hat{\lambda}_{N_0}$

  - Iterate until total number of thresholds is reduced to N-1:
  - Finish: return $T_1, \ldots, T_{N-1}$ and $\hat{\lambda}_0, \hat{\lambda}_1, \ldots, \hat{\lambda}_N$

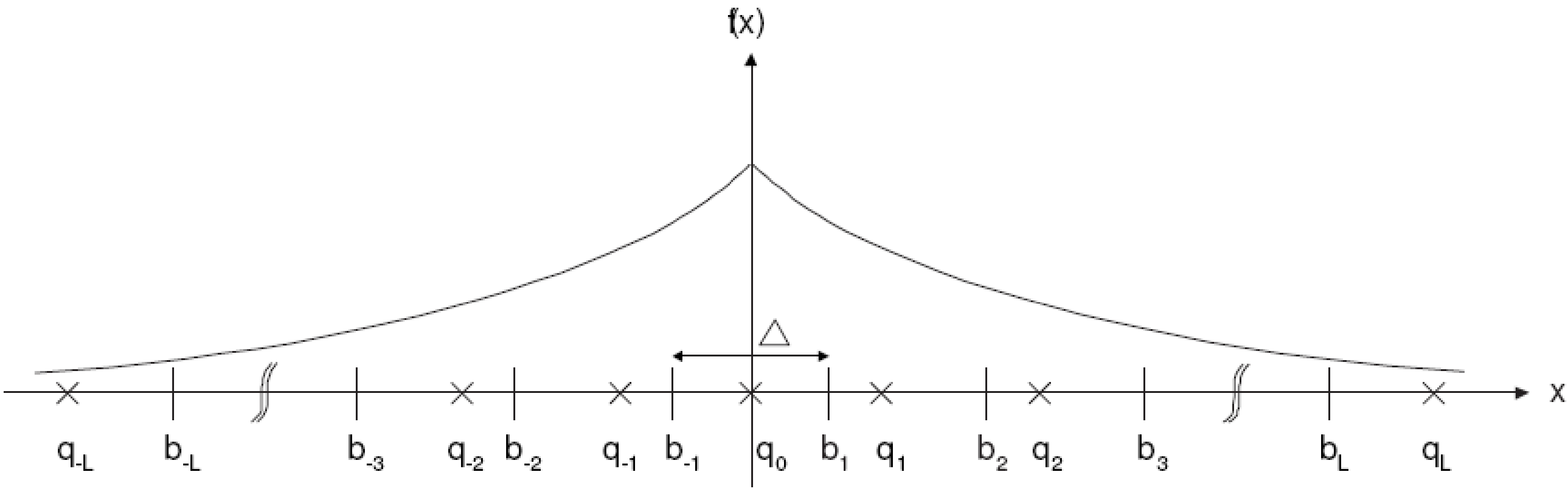
- Side information: Total of 2NS numbers (S=#subbands)

Context-Based MD Wavelet Image Coding: Model-Based Adaptive Quantization

- Purpose: Efficient quantization using a set of quantizers, each customized to an individual class
- Customization is based on the Laplacian parameter estimates, obtained during classification thresholds design
- Two types of quantizers are examined:
  - Uniform Threshold Quantizer (UTQ)
  - Uniform Reconstruction with Unity Ratio Quantizer (URURQ)
- Both types of quantizers well approximate the optimum ECSQ for the Laplacian distribution (with MSE distortion)
- Both are completely defined by a single parameter Δ
- Number of quantization levels is odd ("mid-tread") in order to enable an output entropy < 1 bit/sample

Context-Based MD Wavelet Image Coding: Uniform Threshold Quantizer (UTQ)

- Completely defined by its step size Δ
- Reconstruction levels are optimized for minimum distortion (centroid condition)



Context-Based MD Wavelet Image Coding: Design Strategy for the Quantizers (UTQs)

- Purpose: Avoid complex entropy-constrained design algorithms for the UTQs
- Means: Optimal bit allocation scheme based on a pre-designed array of MSE-optimized UTQs of different step sizes (with no constraint on output entropy)

⇒ Goal: Design an MSE-optimal UTQ with step size $\triangle$ for the Laplacian distribution with parameter λ

  - Expressions for bin boundaries, reconstruction levels and bin probabilities are derived straightforwardly (also found in the literature)

Context-Based MD Wavelet Image Coding: Quantizer Function of UTQ for Laplacian Distribution

- Purpose: Estimate rate and distortion of UTQ to obtain its operational DR function (quantizer function)
  - Quantizer function is required for bit allocation
- Rate *R* is estimated by the output entropy of UTQ:

$$H_Q = -\sum_{j=-L}^{j=L} p_j \log_2 p_j$$

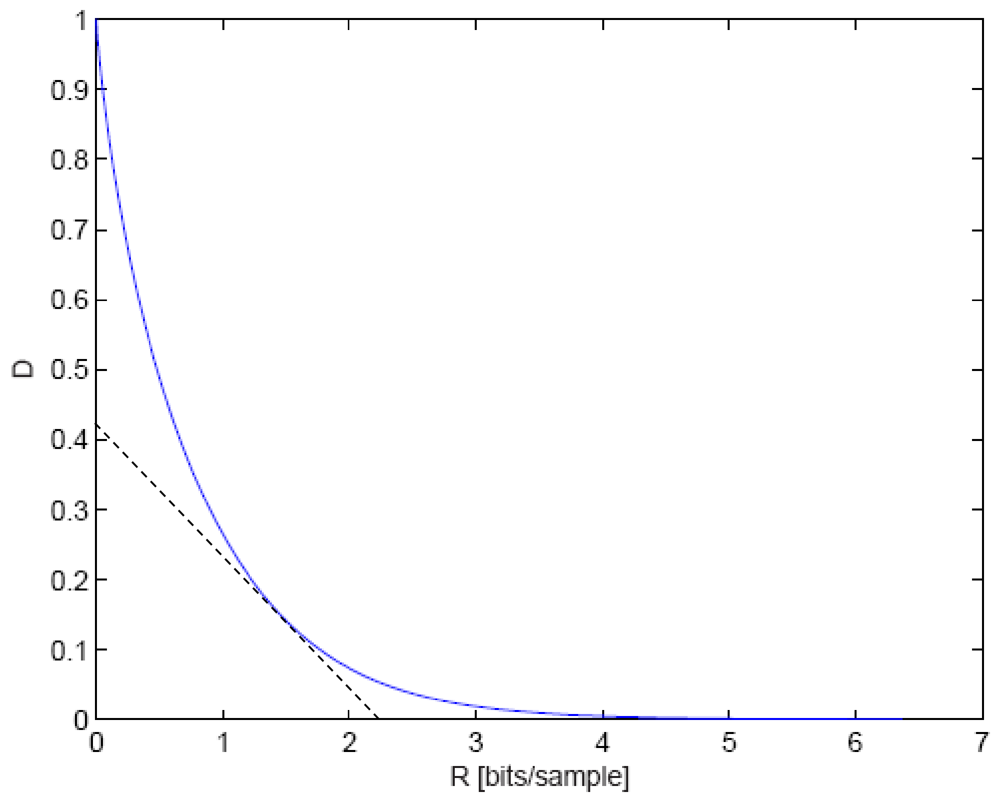
- We derive a closed form expression for the distortion *D* :

$$\begin{aligned} D &= \frac{2}{\lambda^2} - e^{-\lambda\frac{\Delta}{2}}\left(\left(\frac{\Delta}{2}\right)^2 + \frac{\Delta}{\lambda} + \frac{2}{\lambda^2}\right) + \\ &+ \left(\sum_{j=1}^{L-1} e^{-\lambda q_j}\right)\cdot\left[e^{\lambda\delta}\left(\delta^2 - \frac{2\delta}{\lambda} + \frac{2}{\lambda^2}\right) - e^{\lambda(\delta-\Delta)}\left((\delta-\Delta)^2 - \frac{2(\delta-\Delta)}{\lambda} + \frac{2}{\lambda^2}\right)\right] + \\ &+ \frac{1}{\lambda^2}e^{1-\lambda q_L} \end{aligned}$$

where $\delta = \frac{1}{\lambda} - \frac{\Delta}{e^{\lambda\Delta} - 1}$

Context-Based MD Wavelet Image Coding: Quantizer Function of UTQ for Laplacian Distribution (cont.)

- Off-line computation: Array of UTQs obtained for closely spaced values of the step size Δ (for the unit-variance Laplacian distribution)
- Result: Indexed operational DR function (indexed by slope)



Context-Based MD Wavelet Image Coding: Optimal Model-Based Bit Allocation

- Purpose:
  - Given the desired redundancy rate, determine the rate at which each UTQ operates
  - Avoid complex on-line bit allocation procedures
- Means:
  - Efficient optimal bit allocation procedure based on variance scaling and on a pre-designed indexed array of optimized UTQs (fixed resource of the coder, used by encoder and decoder)
  - Performed simultaneously over all classes from all subbands

Context-Based MD Wavelet Image Coding: Optimal Model-Based Bit Allocation (cont.)

- Optimization problem:
  Find the optimal rates $\{R_b\}_{b=1}^{B}$ such that the overall distortion D is minimized, subject to the constraint $R \leq R_T$
- Lagrangian optimization:
  Minimize the Lagrangian cost function $J(\xi) = D + \xi R$ (for a fixed Lagrange multiplier ξ, to be determined such that $R = R_T$)
- Resulting rate allocation equations (≈ "constant slope" principle):

$$D_b'(R_b) = -\frac{\xi}{G_b}, \;\; b = 1, \ldots, B$$

(G_b is the synthesis energy gain factor of class *b*'s subband)

Context-Based MD Wavelet Image Coding: Optimal Model-Based Bit Allocation (cont.)

- Solving the rate allocation equations via variance scaling:

- Observations:
  - For a given rate, the optimal UTQ for input with variance σ^2 is a scaled version, by σ, of the optimal UTQ for unit-variance input
  - The distortion attained by the scaled UTQ above is larger by a factor of σ^2 compared to that attained by the unit-variance UTQ
- Consequence:
  The operational DR function of the optimal UTQ for an input with variance σ^2 is obtained from that for a unit-variance input by scaling the distortion axis by σ^2 (the slope is affected similarly)

⇒ Strategy for solving the rate allocation equation of each class *b* :
  - Slope normalization: $-\xi/G_b \rightarrow -\xi/(G_b\hat{\sigma}_b^2)$ (where $\hat{\sigma}_b^2 = 2/\hat{\lambda}_b^2$)
  - Index the array of UTQs by the normalized slope
  - Scale the obtained "normalized" UTQ by $\hat{\sigma}_b$ to get the actual UTQ

Context-Based MD Wavelet Image Coding: Optimal Model-Based Bit Allocation (cont.)

- Adaptation of the arithmetic entropy encoder:
  - Performed using the bin probabilities of the retrieved UTQ
  - Entropy encoder exploits the higher level statistics captured by the Laplacian model-based classification algorithm
- Determining the Lagrange multiplier ξ such that $R = R_T$:
  - Define:

$$f(\xi) = R^*(\xi) - R_T$$

  - Root finding of $f(\xi)$:
    - Note that $f(\xi)$ is monotonically non-increasing (with ξ)
    - Bracket and use the bisection method (iterate until convergence: $R^* \in (R_T - \epsilon,\ R_T]$)
- The resulting ξ^* is also sent to the decoder

Experimental Results: Configuration

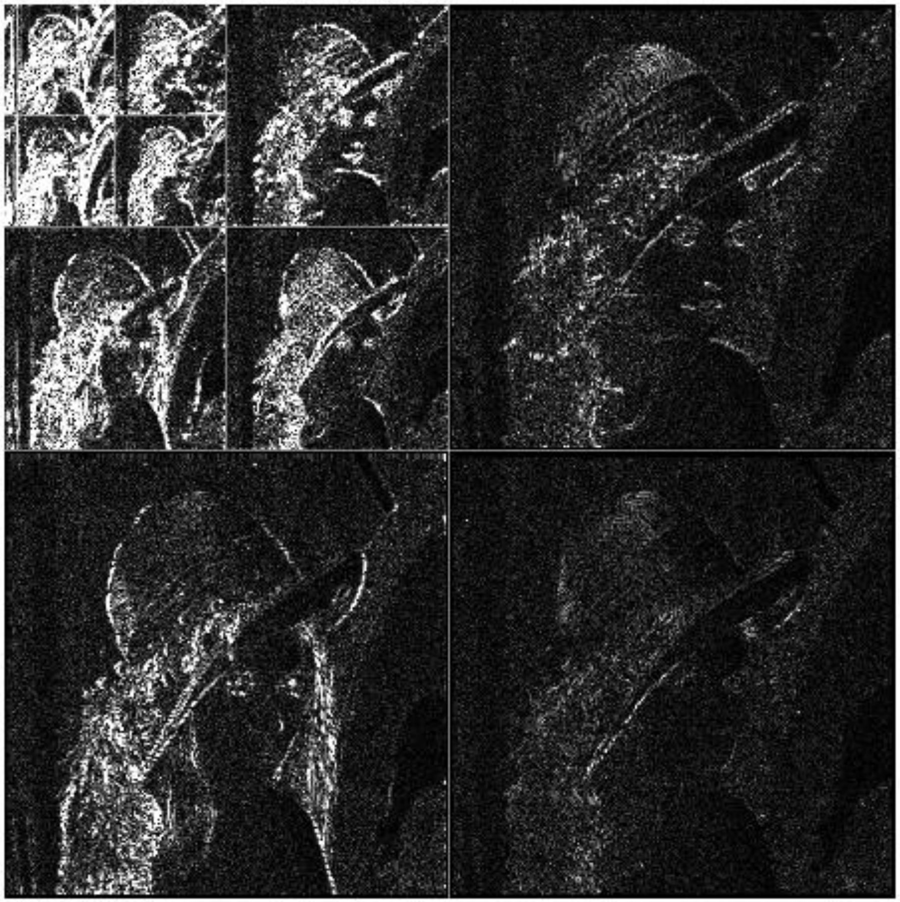
- Balanced case: Descriptions of equal rate and importance

- Wavelet transform:
  - Biorthogonal wavelet transform using Cohen-Daubechies-Feauveau (CDF) 9/7-tap wavelet filters (three-level decomposition)
  - Whole-sample symmetric boundary extension

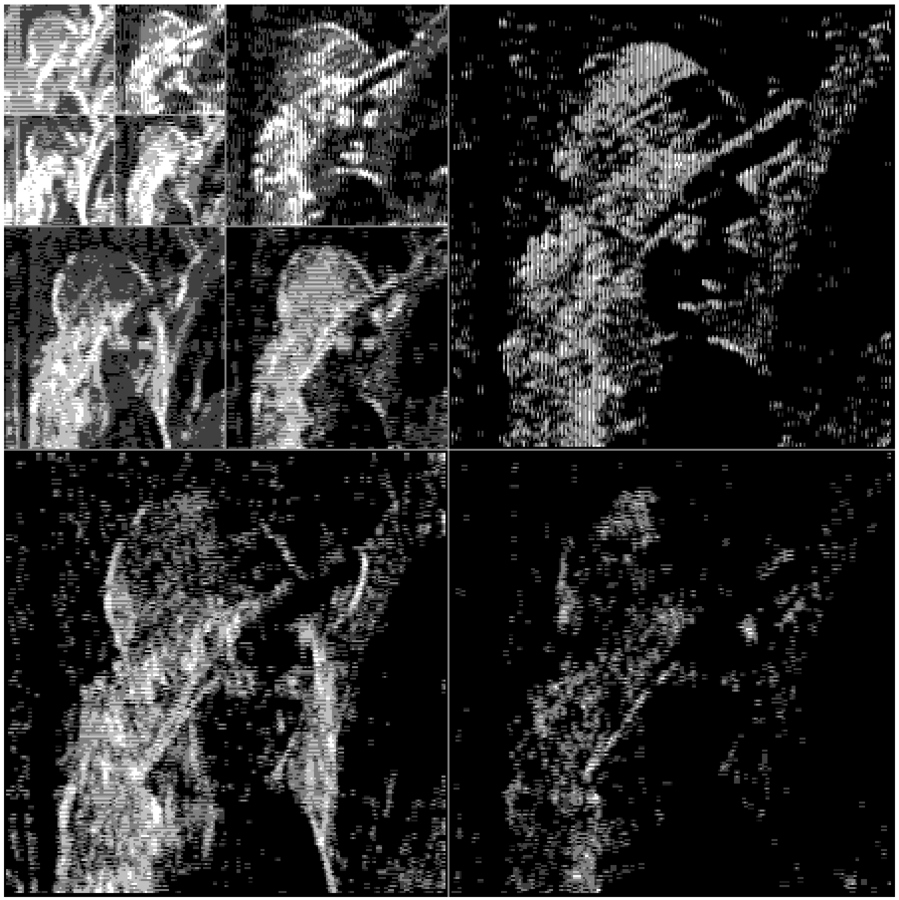
⇒ No coefficient expansion

- Side information parameters are represented using 16 bits (each)
- For demonstrations: Lena image (grayscale), 512x512 pixels (⇒ Overhead rate per desc.: ~0.004 bpp)

Experimental Results: Quality of Classification (Subjective)



Wavelet transform
(differential approximation)

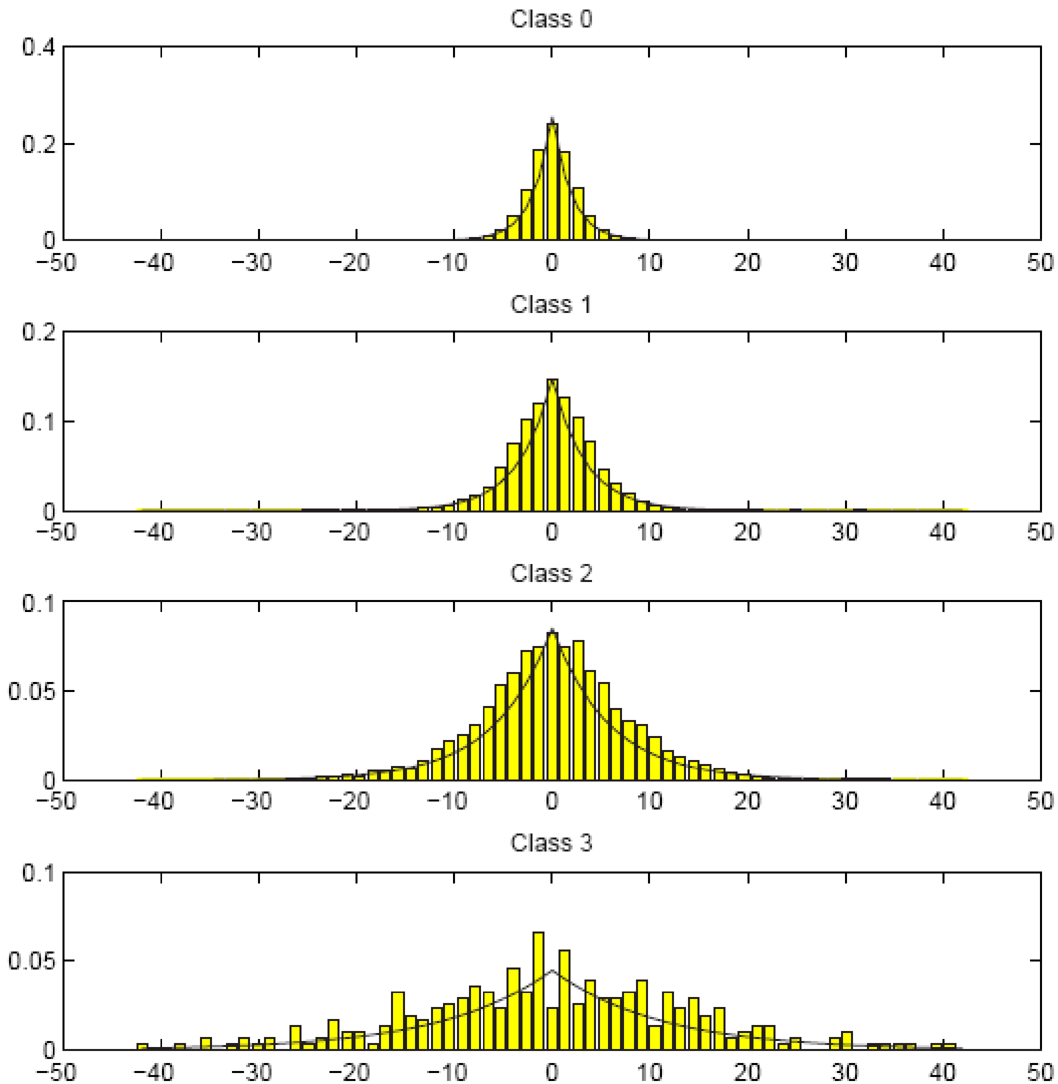


Classification map

Experimental Results: Histograms of Coefficients in the Different Classes

In solid line:

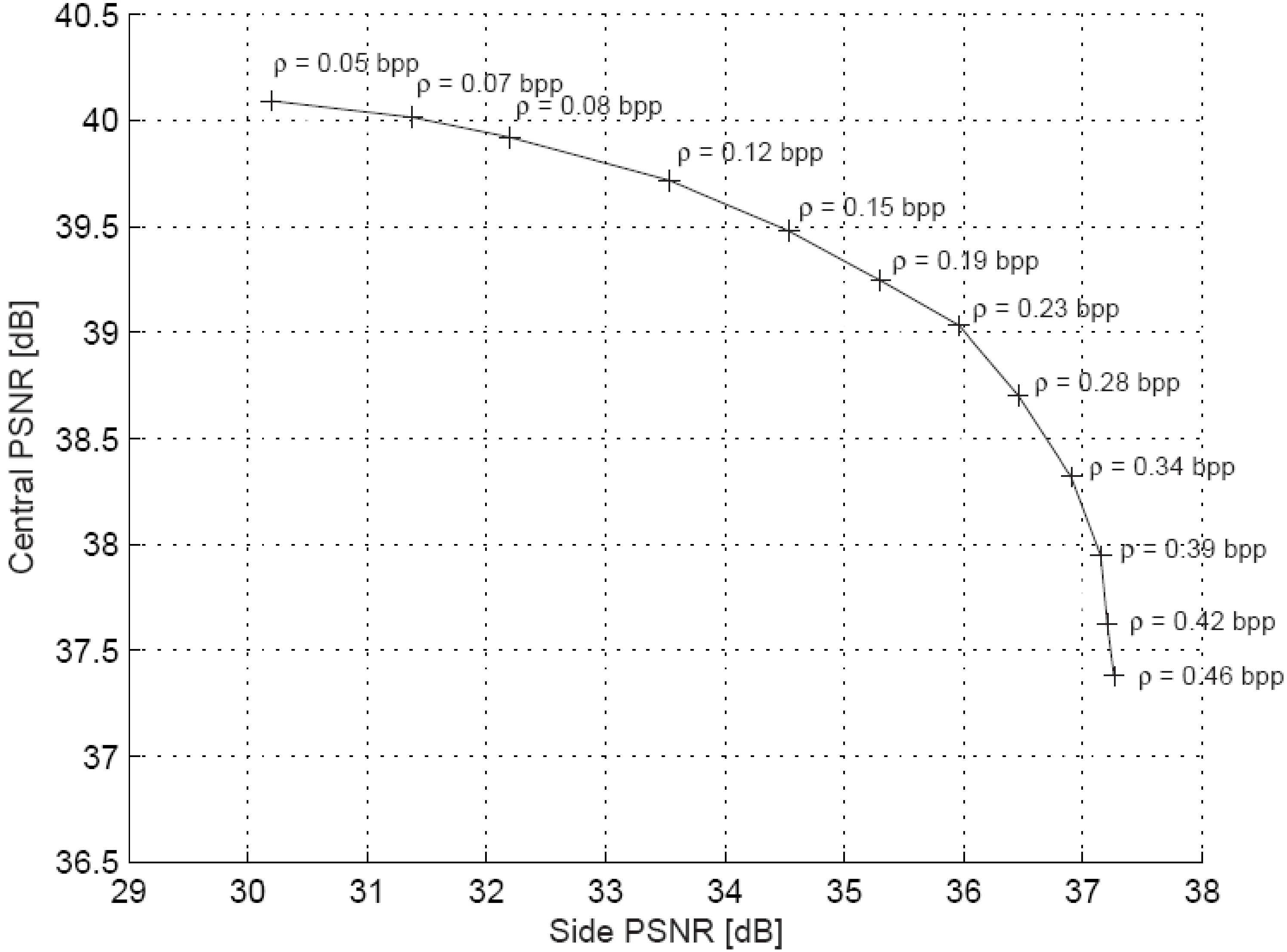
The fitted Laplacian pdf



Experimental Results: Performance of the Proposed Coder

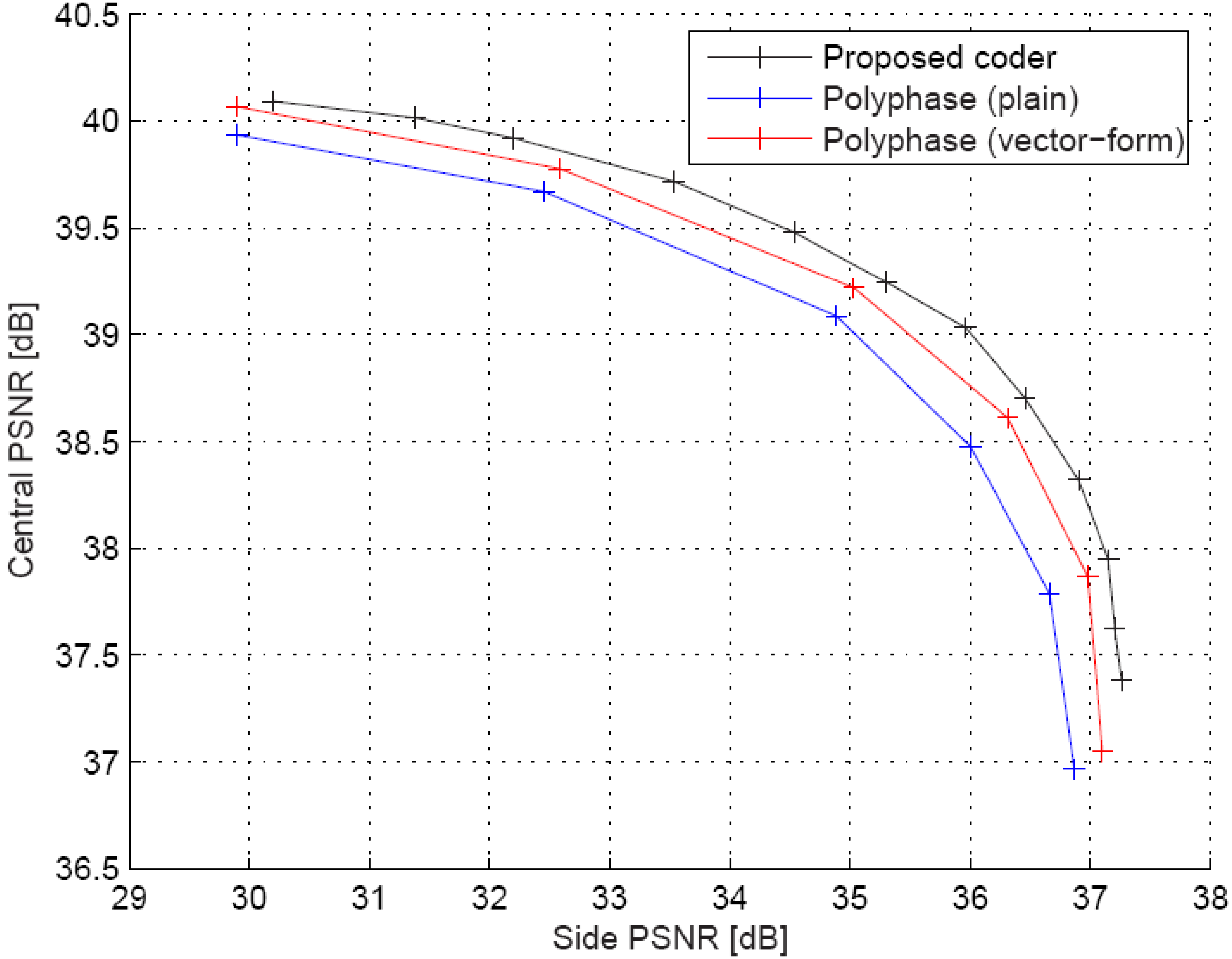
Total rate:

1 bpp



Experimental Results: Performance Compared to Framework Coder

Total rate:
1 bpp



Experimental Results: Subjective Quality of Reconstruction

Original

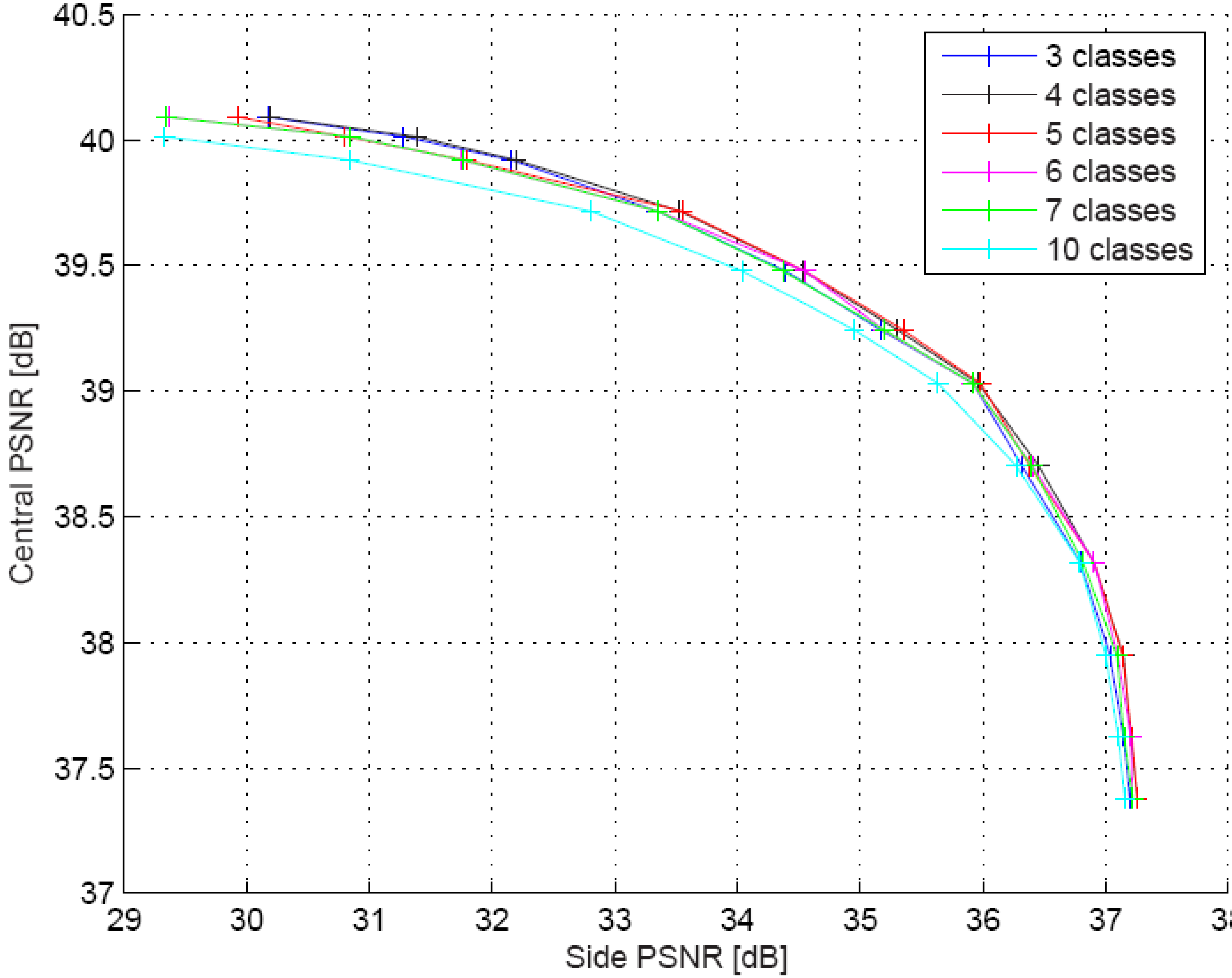
SD
1 bpp
PSNR:
40.63 dB

Proposed
Central rec.
Total: 1 bpp
Red.: 0.46 bpp
PSNR:
37.38 dB

Proposed
Side rec.
Total: 1 bpp
Red.: 0.46 bpp
PSNR:
37.26 dB

Experimental Results: Influence of the Number of Classes on Performance

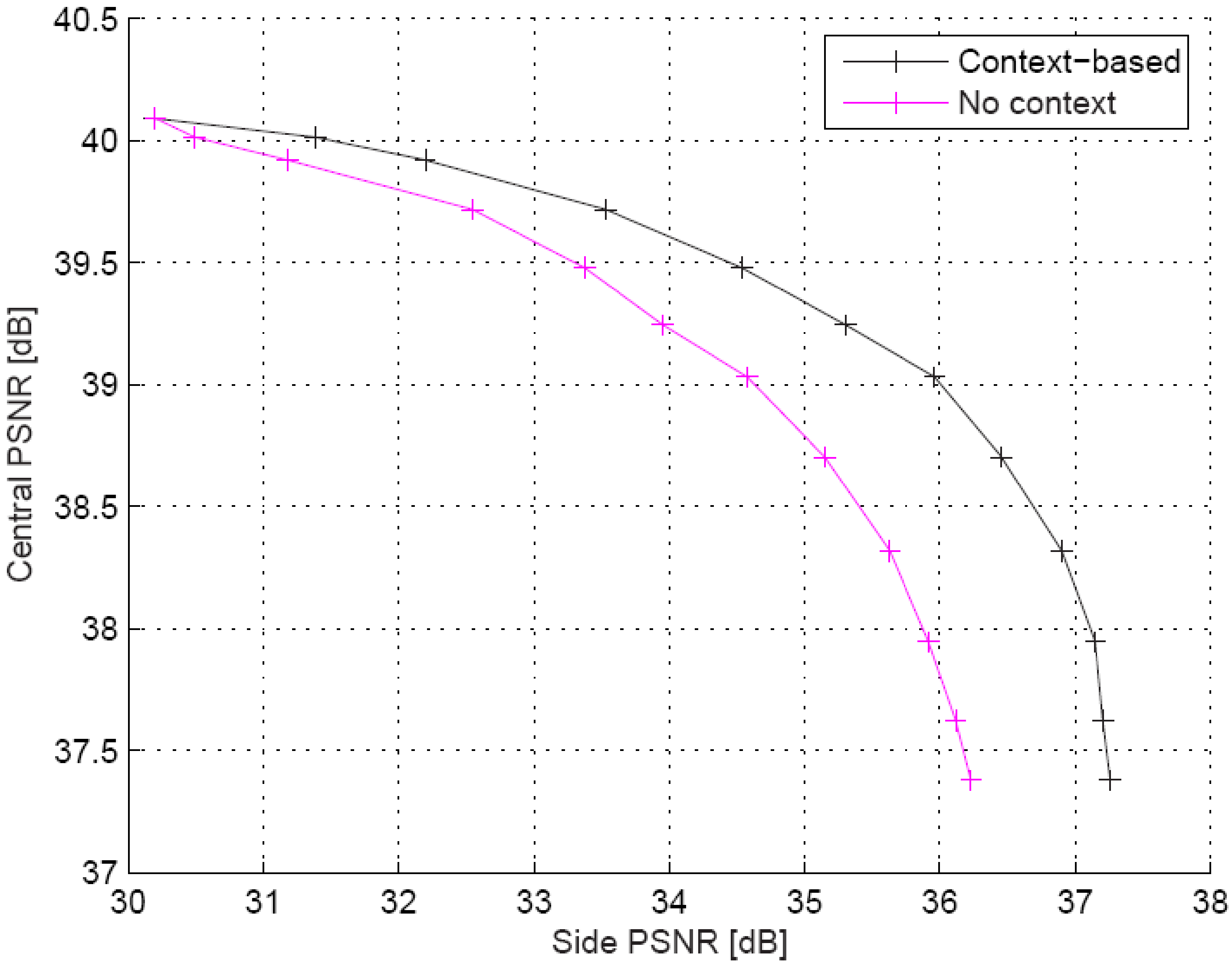
Total rate:
1 bpp



Experimental Results: Context Gain

Total rate:

1 bpp



Experimental Results: Determining Operating Point Based on Channel Properties

- Common channel model:
  - Descriptions are sent over two independent channels
  - Each channel fails with probability *p*
- Problem formulation (for the balanced case):
  - Minimize (subject to a total rate constraint):

$$\bar{D} = (1-p)^2 D_c + 2p(1-p) D_s + p^2 D_{\text{none}}$$

$\Rightarrow$ Minimize (subject to a total rate constraint):

$$D_c + \alpha D_s$$

where $\alpha = \dfrac{2p}{(1-p)}$

Experimental Results: Determining Operating Point Based on Channel Properties (cont.)

- For a fixed total rate (1 bpp):

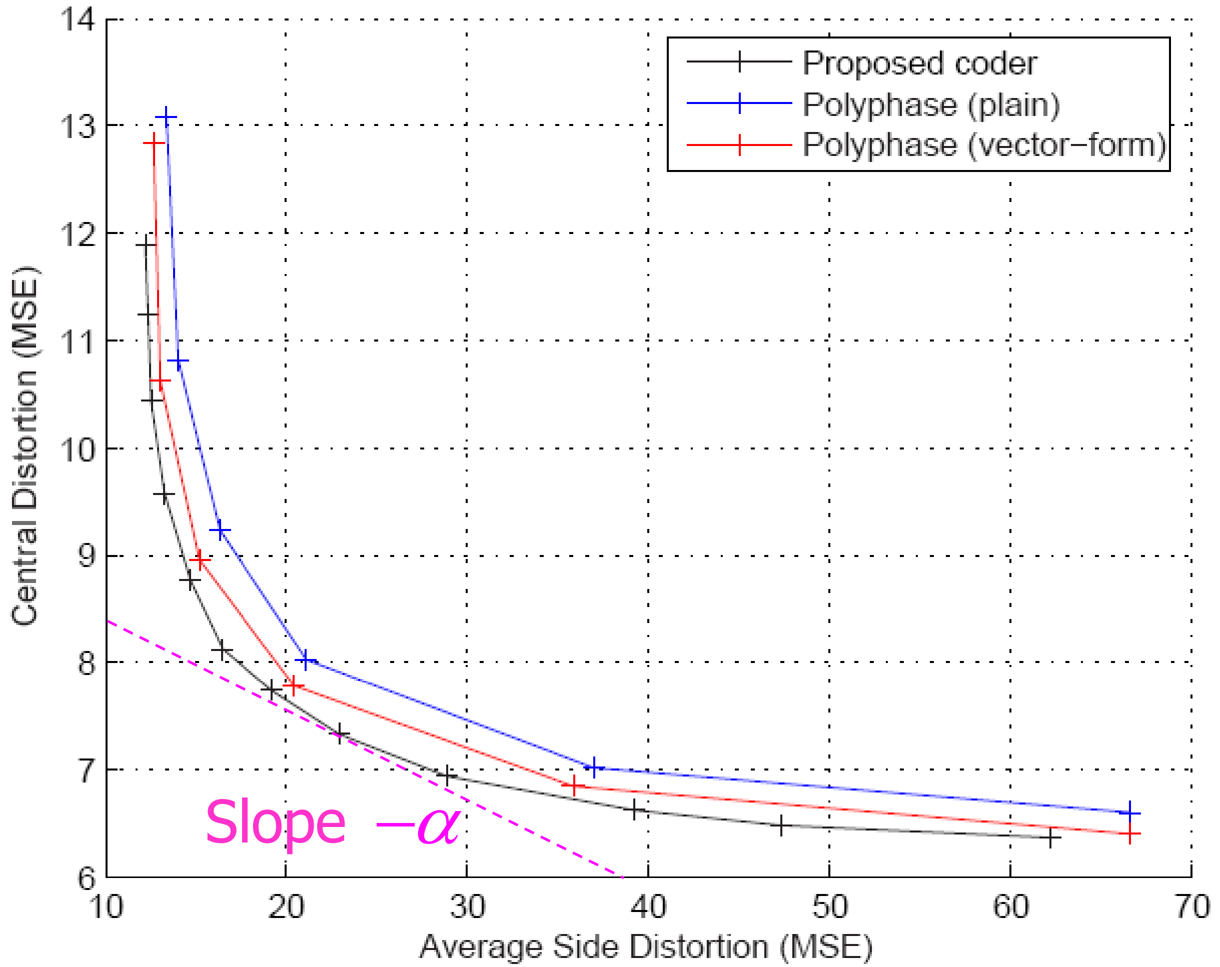
min { $D_c + \alpha D_s$ }

Note:

$p \rightarrow 0^+ \Rightarrow \alpha \rightarrow 0$

$p \rightarrow 1^- \Rightarrow \alpha \rightarrow \infty$

Finding the optimum: 1-D minimization (e.g., using golden section search)

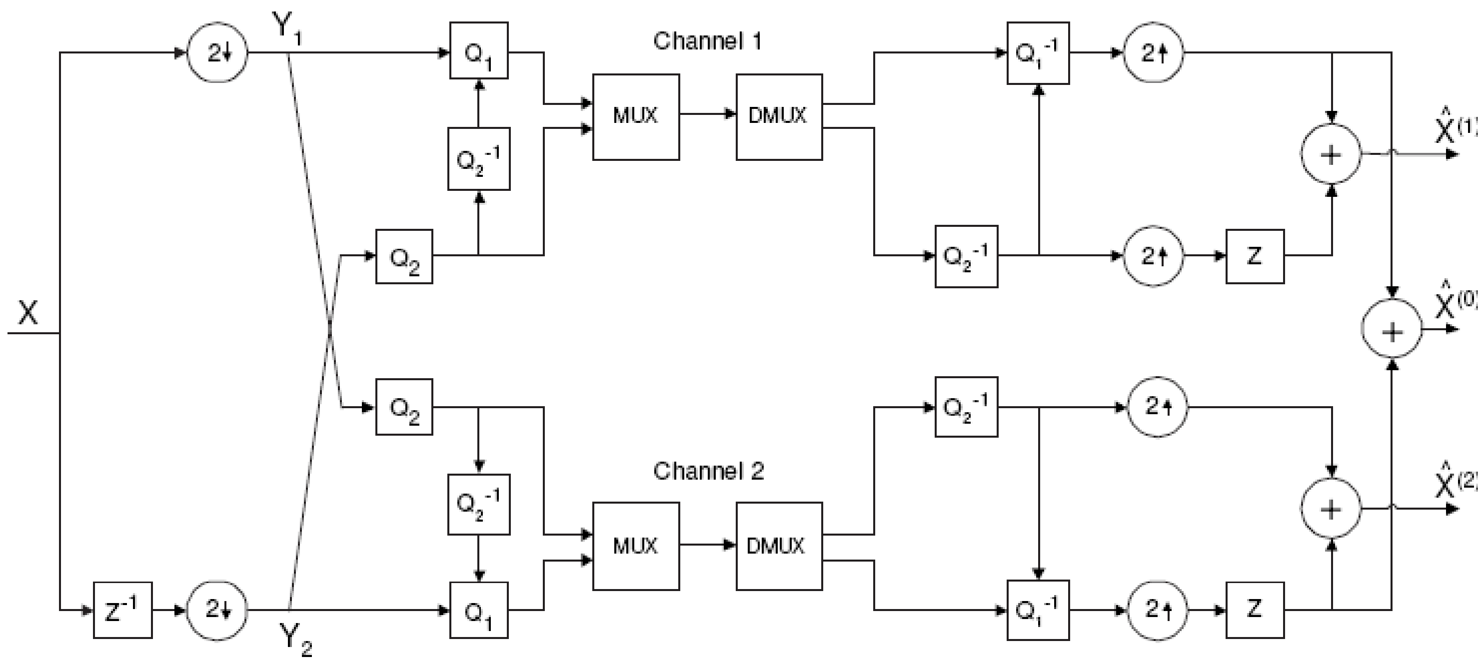


Summary

- Introduction and fundamentals of MD coding
- Framework: MD coding via polyphase transform
- Proposed context-based MD wavelet image coder
  - Context formation
  - Context-based classficiation
  - Model-based adaptive quantization
  - Optimal model-based bit allocation
- Experimental results
  - Classification results
  - RD performance (also compared to framework)
  - Context gain
  - Determining the optimal operating point (for a given channel)

Future Directions

- Reverse context-based MD coding system:



- Utilization of across-scale dependencies
- Extensions to more than two descriptions

References

- [El Gamal and Cover, 1982]

  A. A. El Gamal and T. M. Cover, “Achievable rates for multiple descriptions,” *IEEE Trans. Inf. Theory*, vol. 28, no. 6, pp. 851–857, Nov 1982.

- [Goyal, 2001]

  V. K. Goyal, “Multiple description coding: compression meets the network,” *IEEE Signal Process. Mag.*, vol. 18, no. 5, pp. 74–93, Sep 2001.

- [Jiang and Ortega, 1999]

  W. Jiang and A. Ortega, “Multiple description coding via polyphase transform and selective quantization,” in *Proc. Visual Communic. and Image Proc.*, San Jose, CA, Jan 1999, pp. 998–1008.

References (cont.)

[Buccigrossi and Simoncelli, 1999]

R. W. Buccigrossi and E. P. Simoncelli, "Image compression via joint statistical characterization in the wavelet domain," *IEEE Trans. Image Process.*, vol. 8, no. 12, pp. 1688–1701, Dec 1999.

- [Yoo et al., 1999]

Y. Yoo, A. Ortega, and B. Yu, "Image subband coding using context-based classification and adaptive quantization," *IEEE Trans. Image Process.*, vol. 8, no. 12, pp. 1702–1715, Dec 1999.

Thank You!