CSEP 590
Data Compression
Autumn 2007

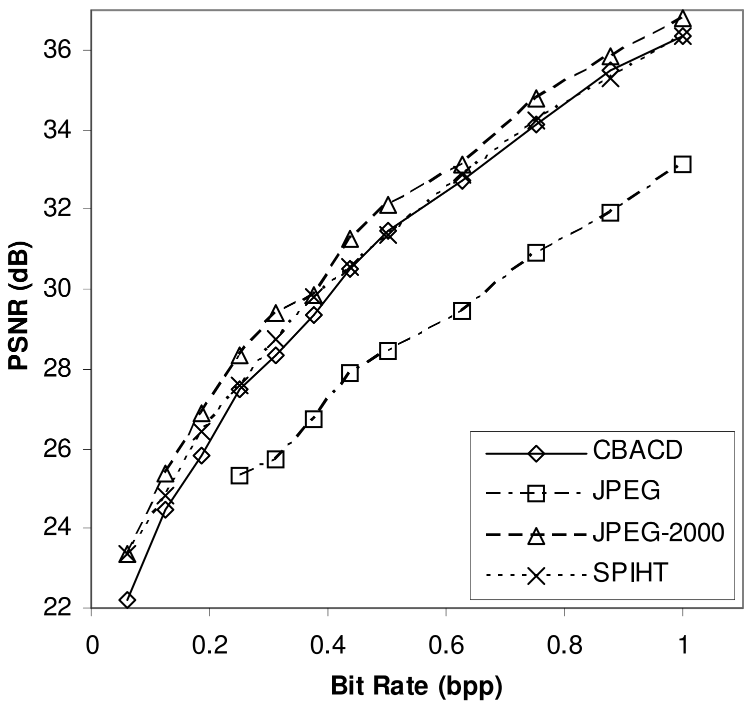
Context Based Arithmetic Coding for the DCT (CBACD)

(Kyle Littlefield, 2006)

CBACD overview

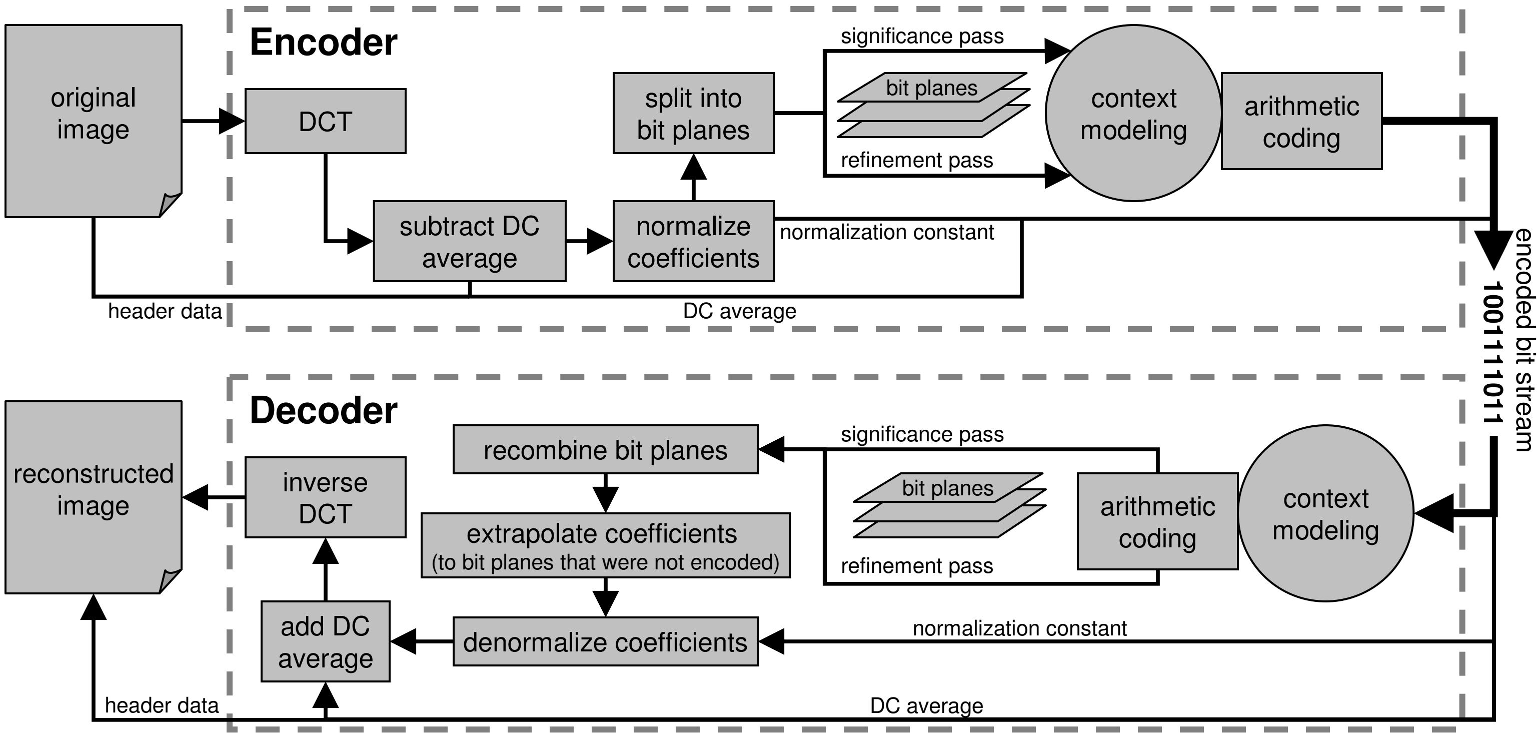
- Evolved out of PACW
  - A simple wavelet based coder developed at UW by Dane Barney and Amanda Askew
- Goals:
  - Replace wavelet transform with the DCT
  - Replace context model with one suitable for the DCT
  - Compare performance to
    - Existing DCT-based methods, primarily JPEG
    - State of the art wavelet methods

CBACD Overview - Results



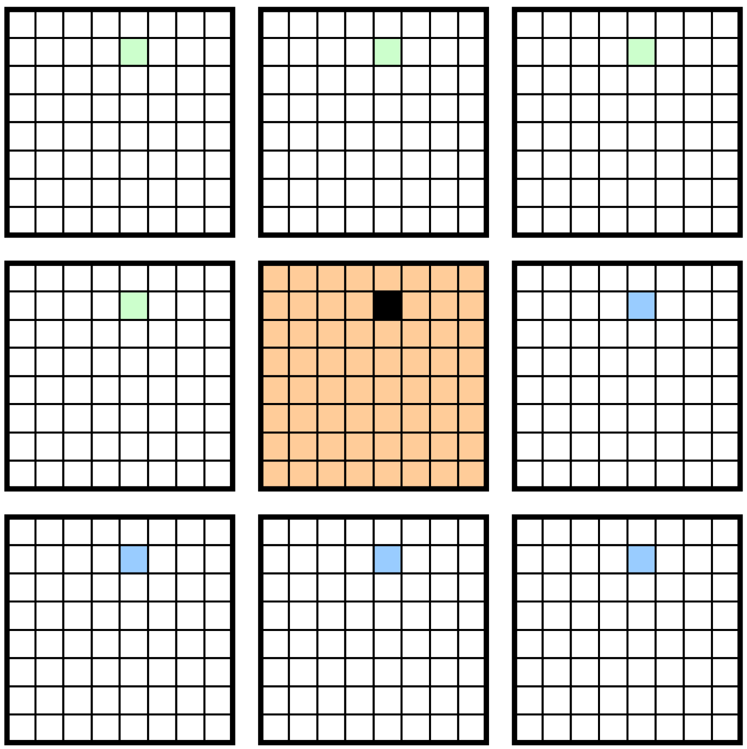
- Performs significantly better than JPEG
- Performs slightly under wavelet based methods such as SPIHT and JPEG-2000

CBACD Overview



Context Modeling – Significance Bits

- Based on two factors
  - Intra-block correlation: Relationships between subbands within a block.
  - Inter-block correlation: Relationships to neighboring blocks, within the same subband.



- ■ Coefficient of interest
- Intra-block/inter-subband coefficients
- Inter-block/intra-subband coefficients (For which information from the current bit-plane is available.)
- Inter-block/intra-subband coefficients (For which only information from the previous bit-plane is available.)

Context Modeling – Significance Bits

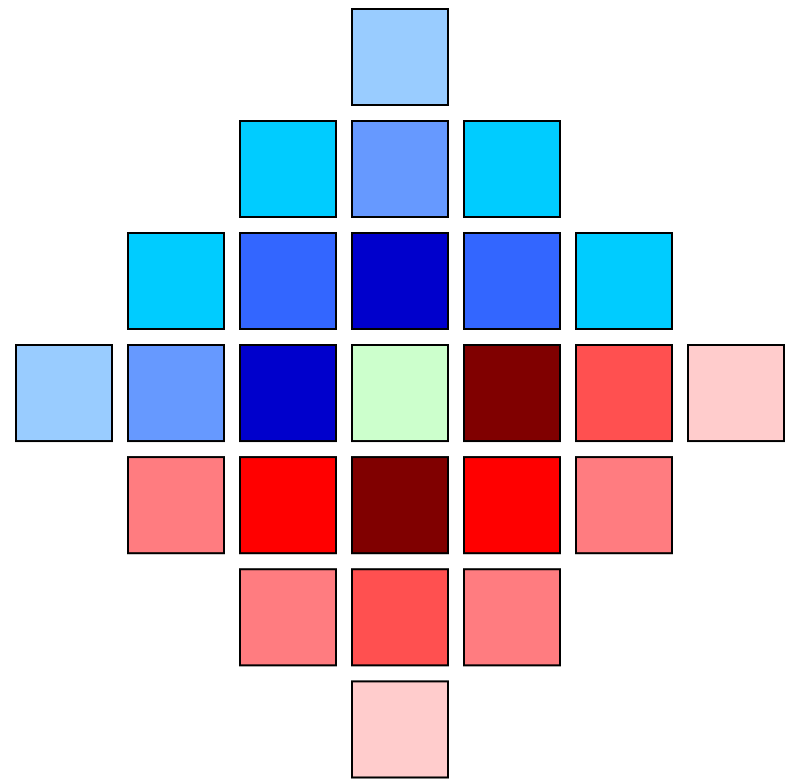
- First a significance factor is computed, based on a linear sum of the two factors

$$f(sub,x,y) = c_{spatial} \sum_{i,j} \frac{isSig(sub, x+i, y+j)}{dist(i,j)} + c_{frequency} \sum_{i=1}^{63} isSig(i,x,y) + c_{constant}$$

- Details
  - $c_{spatial}$ is set to 1.5, $c_{frequency}$ to 0.225, $c_{constant}$ to .375
  - The distance formula is taken to be the square of the Euclidean distance: i^2+j^2

Context Modeling – Significance Bits

- Interblock correlation sum is taken over 24 surrounding blocks. Closer blocks have more influence.
- Blue denotes blocks for which information is available for the current bit plane.
- Red denotes blocks for which information is only available for the previous bit-plane.



Context Modeling – Significance Bits

- A context is determined from the significance factor by truncating to an integer which is used to look up the context.
- A maximum of five contexts are used per subband (each subband is treated separately).
  - All significance factors larger than 4 are truncated to 4.

Context Dilution Concerns

- Context dilution occurs when only a few bits are encoded in each context
  - Results in decreased arithmetic coding performance, as contexts do not have enough bits encoded to meaningfully update statistics
- Most context-based coders use many fewer contexts than CBACD
  - EBCOT – 27
  - PACW – 30-56 (variable)
  - JPEG-2000 – 27
- CBACD uses 340 contexts

Context Dilution Experiments

- Context dilution concerns were approached by trying various grouping of subbands.
- If context dilution was occuring, grouping subbands should result in large improvements in PSNR

Subband groupings

0	1	2	3	4	5	6	7
8	9	10	11	12	13	14	15
16	17	18	19	20	21	22	23
24	25	26	27	28	29	30	31
32	33	34	35	36	37	38	39
40	41	42	43	44	45	46	47
48	49	50	51	52	53	54	55
56	57	58	59	60	61	62	63

Original CBACD
(64 groups / 320 contexts)

0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Single group
(1 group / 5 contexts)

More subband groupings

0	1	2	3	4	5	6	7
1	1	2	3	4	5	6	7
2	2	2	3	4	5	6	7
3	3	3	4	5	5	6	7
4	4	4	5	5	6	7	8
5	5	5	5	6	7	7	8
6	6	6	6	7	7	8	9
7	7	7	7	8	8	9	9

Circular (10 groups / 50 contexts)

0	1	2	3	4	5	6	7
10	1	2	3	4	5	6	7
11	11	2	3	4	5	6	7
12	12	12	4	5	5	6	7
13	13	13	14	5	6	7	8
14	14	14	14	15	7	7	8
15	15	15	15	16	16	8	9
16	16	16	16	17	17	18	9

Circular with horizontal/vertical split (19 groups / 95 contexts)

More subband groupings

0	1	2	3	4	5	6	7
1	2	3	4	5	6	7	8
2	3	4	5	6	7	8	9
3	4	5	6	7	8	9	10
4	5	6	7	8	9	10	11
5	6	7	8	9	10	11	12
6	7	8	9	10	11	12	13
7	8	9	10	11	12	13	14

Diagonal
(15 groups / 75 contexts)

0	1	2	3	4	5	6	7
1	1	2	3	4	5	6	7
2	2	2	3	4	5	6	7
3	3	3	3	4	5	6	7
4	4	4	4	4	5	6	7
5	5	5	5	5	5	6	7
6	6	6	6	6	6	6	7
7	7	7	7	7	7	7	7

Max frequency
(8 groups / 40 contexts)

More subband groupings

0	1	2	3	4	5	6	7
8	1	2	3	4	5	6	7
9	9	2	3	4	5	6	7
10	10	10	3	4	5	6	7
11	11	11	11	4	5	6	7
12	12	12	12	12	5	6	7
13	13	13	13	13	13	6	7
14	14	14	14	14	14	14	7

Max frequency with horizontal/vertical split (15 groups/75 contexts)

0	1	3	3	3	6	6	6
2	5	5	5	6	6	6	6
4	5	5	5	6	6	6	6
4	5	5	6	6	6	6	6
4	6	6	6	6	6	6	6
6	6	6	6	6	6	6	6
6	6	6	6	6	6	6	6
6	6	6	6	6	6	6	6

Generic grouping by types (7 groups / 35 contexts)

Context Dilution Results

- Over a set of six images encoded at 0.25 bits per pixel (changes in PSNR dB)
  - Single grouping: -0.128
  - Circular: +0.006
  - Circular with horizontal/vertical split: +0.008
  - Diagonal: +0.016
  - Max Frequency: +0.002
  - Max Frequency with horizontal/vertical split: +0.006
  - Grouping by type: -0.004

Context Dilution - Conclusions

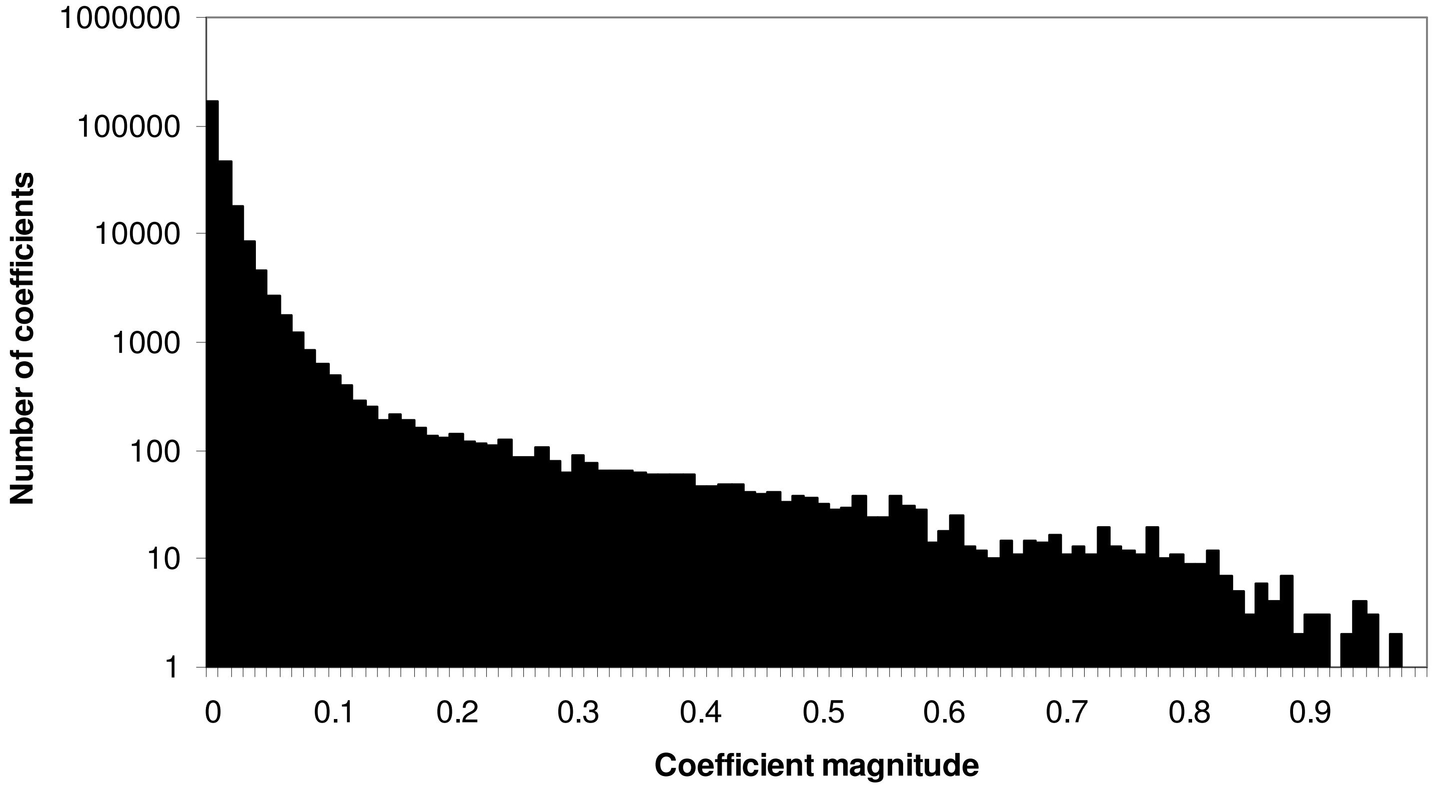
- The single grouping (not surprisingly) does significantly worse than any other
- The other groupings perform about the same
  - The original CBACD is among the worst of these
- Context dilution is only a very slight concern within the CBACD architecture

Context Modeling – Sign Bits

- Modeled similar to refinement bits, except:
  - No intra-block correlation
  - Smaller area over which sum is taken
  - All subbands use a single set of contexts
  - 9 contexts used (instead of 5)

Context Modeling – Refinement Bits

- Model is based on coefficient distribution
  - Coefficients are skewed towards 0, so refinement bits are also skewed towards 0



Context Modeling – Refinement Bits

- Bits are placed into contexts based on the number of bit planes since the coefficient became significant
- Provides significant improvements over putting all refinement bits in 1 context
  - Leads to PSNR improvements of up to .1 dB at high bit rates

	# 0s	# 1s	% 0s
1	26758	14575	64.74
2	13737	9511	59.09
3	3610	5324	54.23
4	2927	2677	52.23
5	1491	1436	50.93
6	649	650	49.96

Coefficient Extrapolation

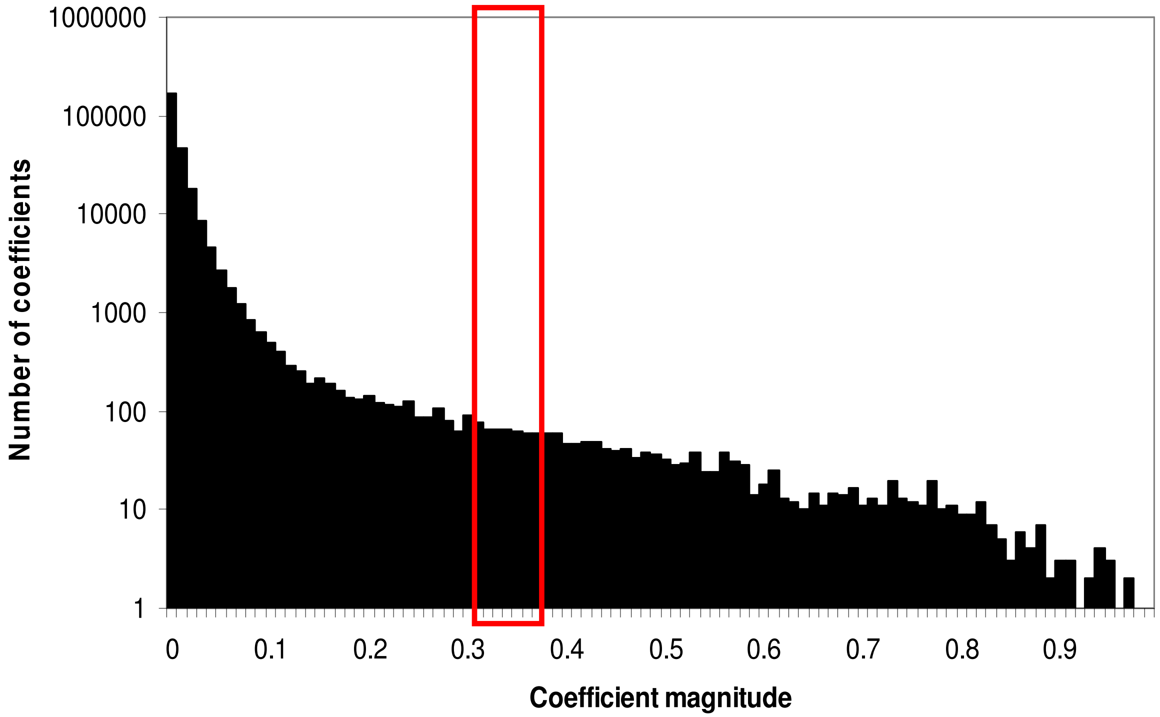
- The bit plane process guarantees that only the first few bits of each coefficient will be transmitted to the decoder.
  - The decoder must decide how to fill in the untransmitted bits

Transmitted coefficient	.0101????????...
Possible extrapolations	.010100000000… .010111111111… .010110000000… .010101011001…

Coefficient Extrapolation - Goal

- Reduce reconstructed image distortion
- Reduce distortion as measured by PSNR
- Reduce average per-pixel difference between original and deconstructed image
- Reduce average difference between original DCT coefficient and reconstructed DCT coefficient
- Reduce average distance across possible DCT coefficient distribution.

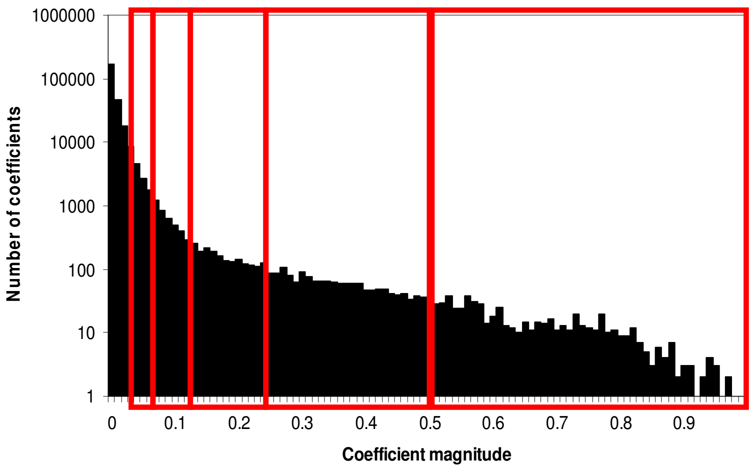
Coefficient Extrapolation Example



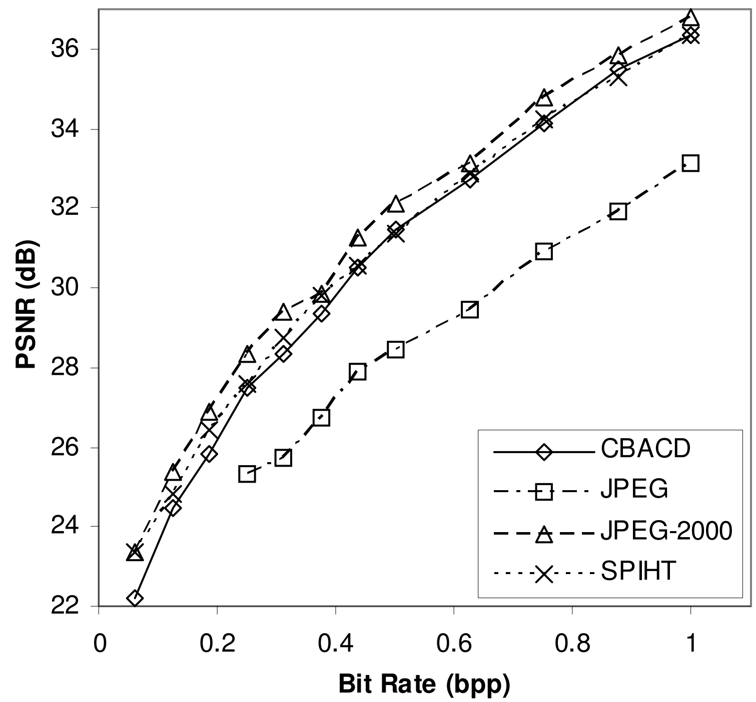
Best Extrapolation for .0101????????... is .010101101100…

Coefficient Extrapolation

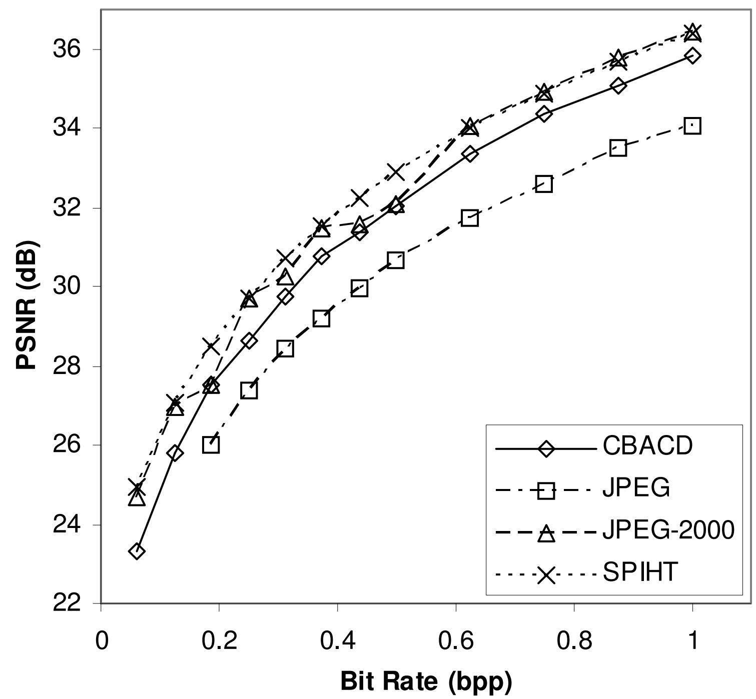
- Can not calculate best extrapolation for every possible set of transmitted bits
- Instead, compute best extrapolation for every possible pair of
  - Bit plane that the coefficient became significant in
  - Number of bits transmitted after the coefficient became significant
- Best extrapolation is computed separately for each subband
- The 'standard' coefficient distribution consists of the sum of the distributions from 250 images pulled from the CBIRT database (http://www.cs.washington.edu/research/imagedatabase/groundtruth/)



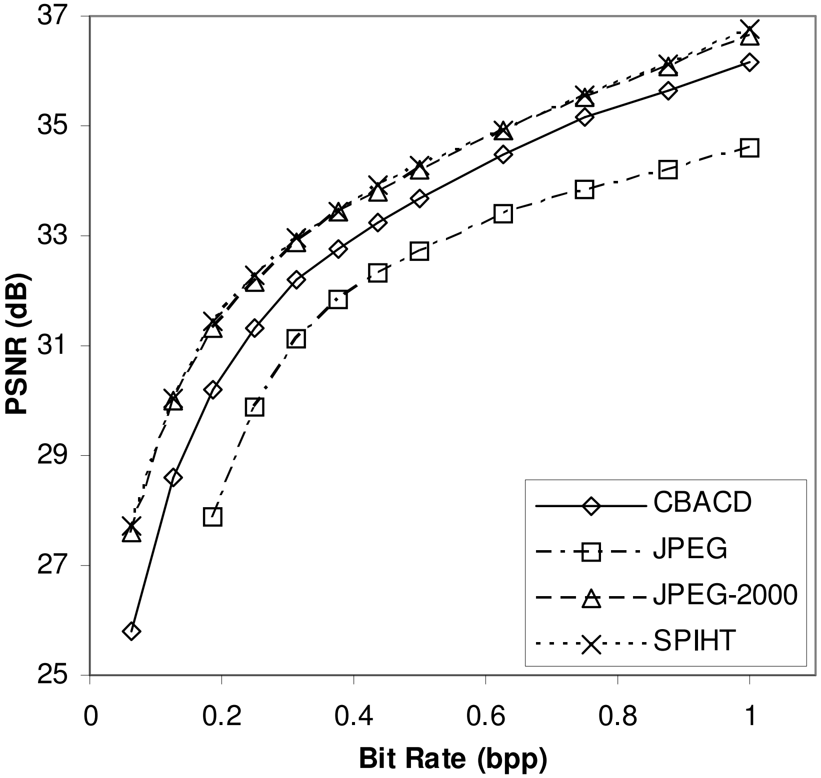
Results - Barbara



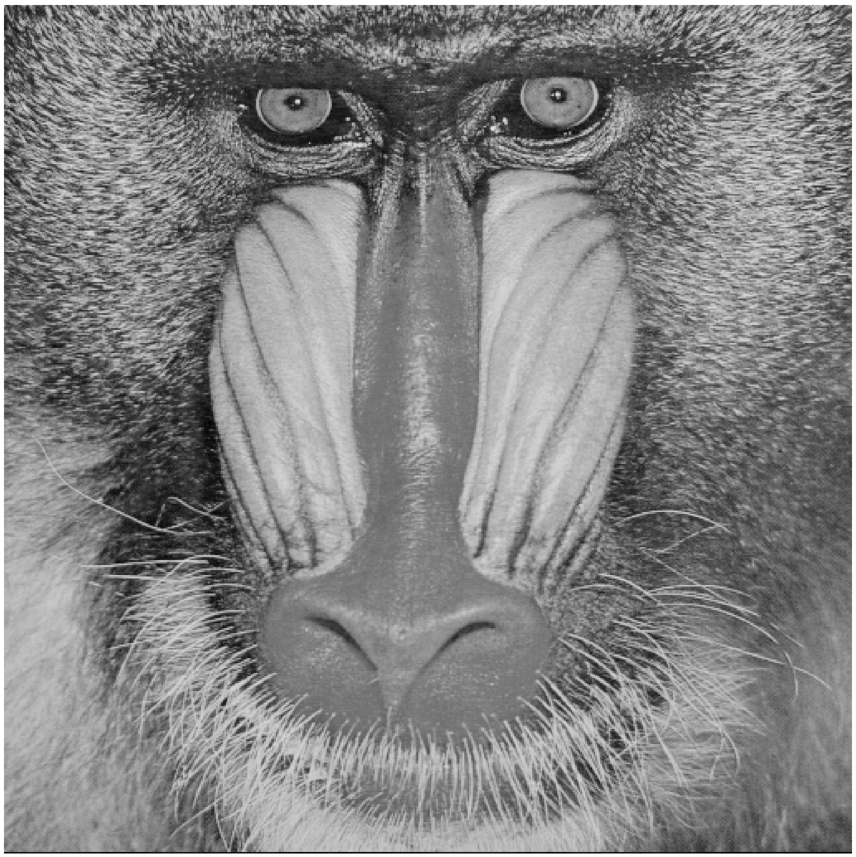
Results - Boat

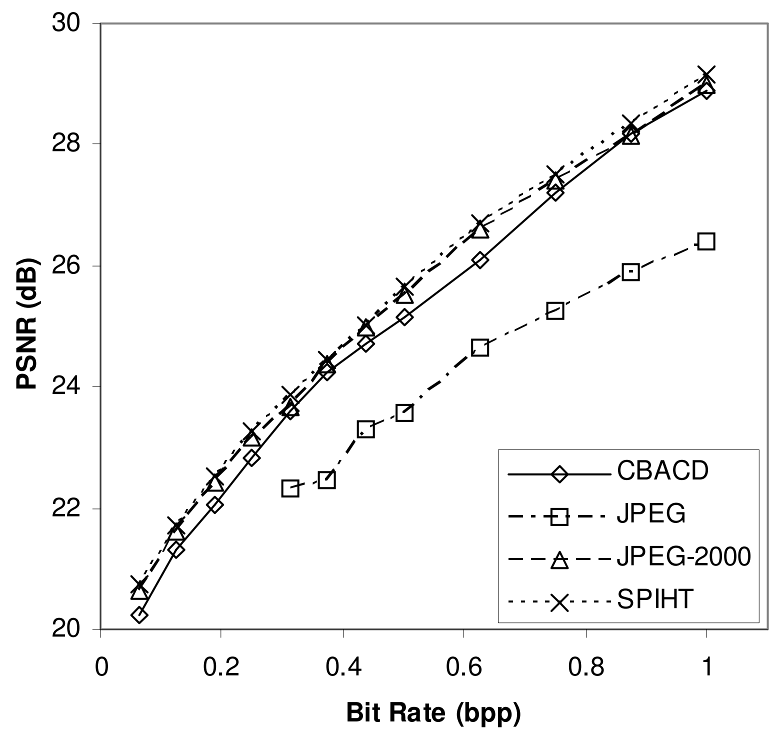


Results – Lena

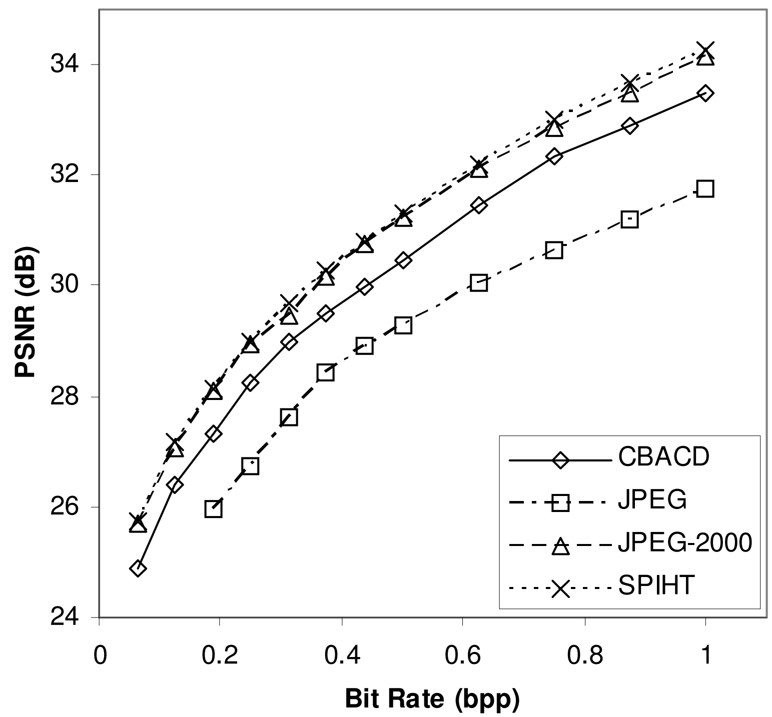


Results - Mandrill

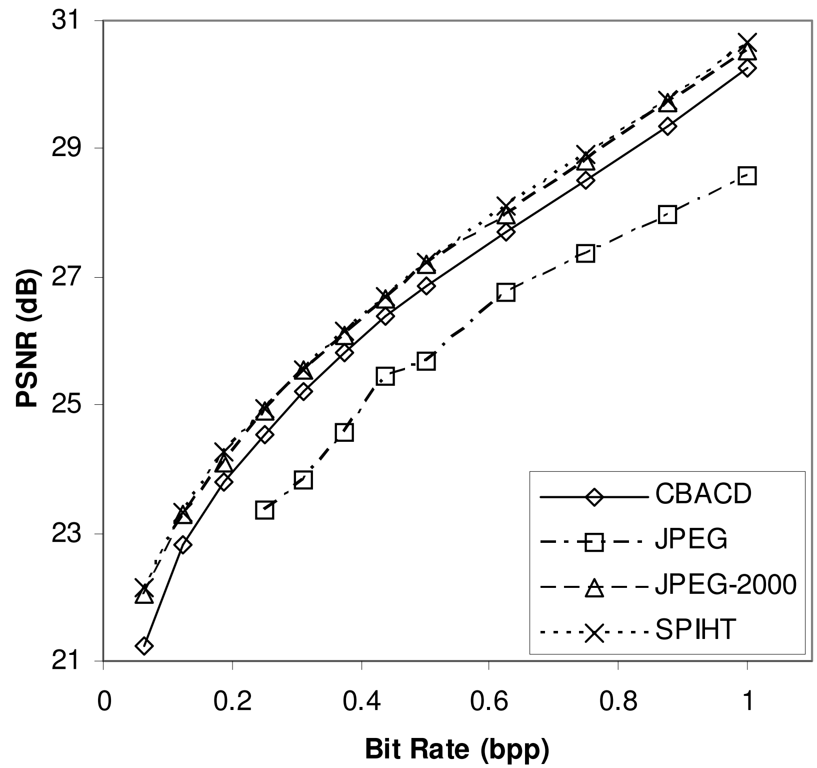




Results - Pentagon



Results - Stream



Results

- For images involving a lot of high frequency information, such as Barbara and Mandrill
  - CBACD performance is a fraction of a dB worse than wavelet methods.
- For images with fewer high-frequency components, such as Pentagon
  - CBACD performs somewhat worse than wavelet methods, from .5 dB at high bit rates, up to 1.5 dB at very low bit rates.