Basic Image Compression
or
Compression 101

Bernie Brower
Kodak
beranrd.brower@kodak.com

Derived from tutorial from Majid Rabbani

BWC 101 Outline

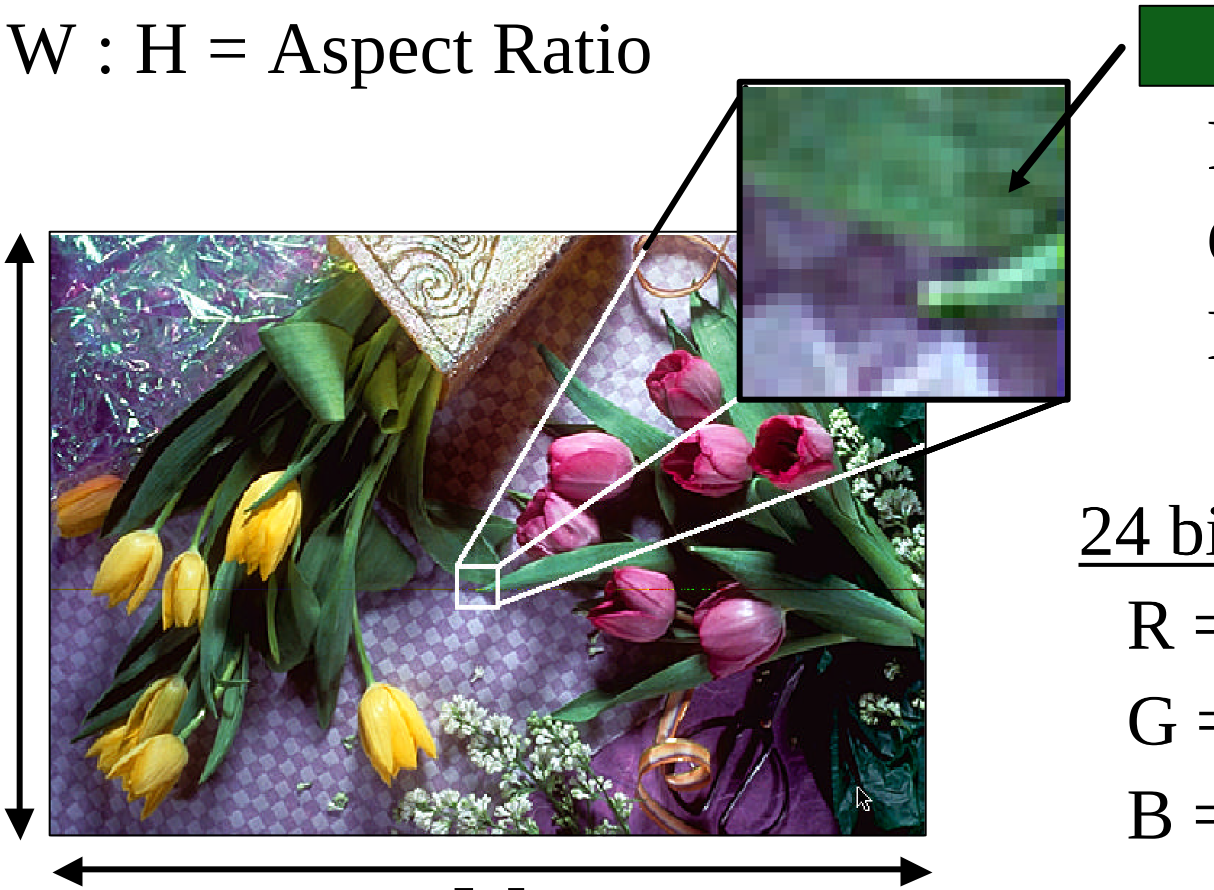
- Motivation for image compression
- Why images can be compressed
- Image compression basics
  - Transform (DPCM, DCT, Wavelet)
  - Quantization
  - Encoding (Huffman, arithmetic)
  - Rate Control
  - Color and Multidimensional transforms
- Selecting a compression algorithm
- Image compression standards
- Image compression systems
  - JPEG DCT
  - JPEG 2000
- Current Dissemination of Imagery

Digital Images

- Digital images are produced by:
  - Remote sensing satellites and aircraft
  - Digitized film and paper
  - Handheld cameras
  - Synthetic imaging (computer generation)
- Regardless, digital images are sampled in;
  - Space ⇨ Resolution (GSD, DPI)
  - Amplitude ⇨ Bit depth (bits per pixel)
  - Spectral ⇨ Resolution (# of bands/color)
- The area coverage or the size of the image "footprint" is dependent on the field of view of the camera system
- [High resolution] x [large bit depth] x [large size] x [large # of bands] = Storage and Transmission challenges

Digital Image Representation

W : H = Aspect Ratio



H

W

Pixel

Red = 15

Green = 95

Blue = 27

24 bits per Pixel

R = 0 → 255

G = 0 → 255

B = 0 → 255

Bits and Bytes

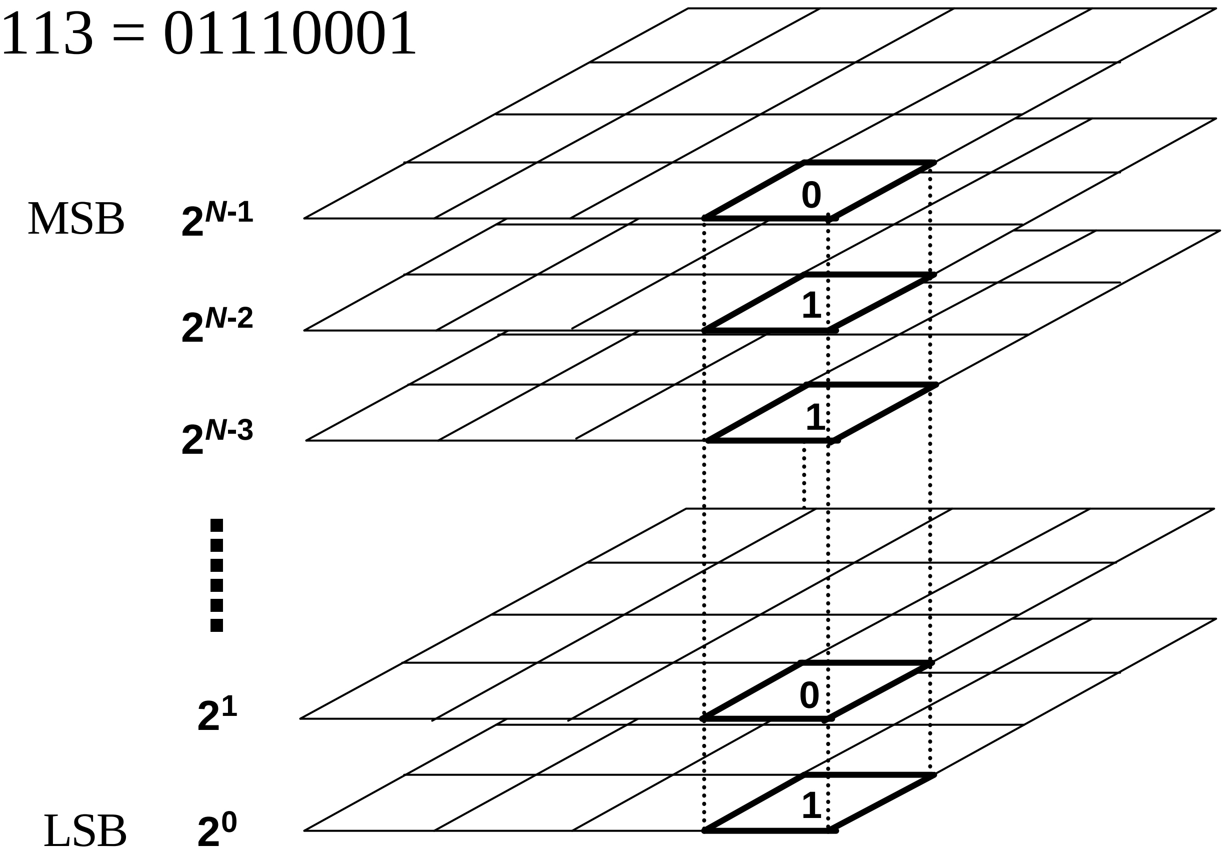
The following terminology is often used when referring to digital image sizes (**bit** is abbreviation for binary digit):

bit (b)	1 bit	1 bit	
byte (B)	8 bits	8 bits	
kilobyte (KB)	2^{10} bytes	10^{3}	+ 2.4%
megabyte (MB)	2^{20} bytes	10^{6}	+ 4.9%
gigabyte (GB)	2^{30} bytes	10^{9}	+ 7.4%
terabyte (TB)	2^{40} bytes	10^{12}	+ 10%

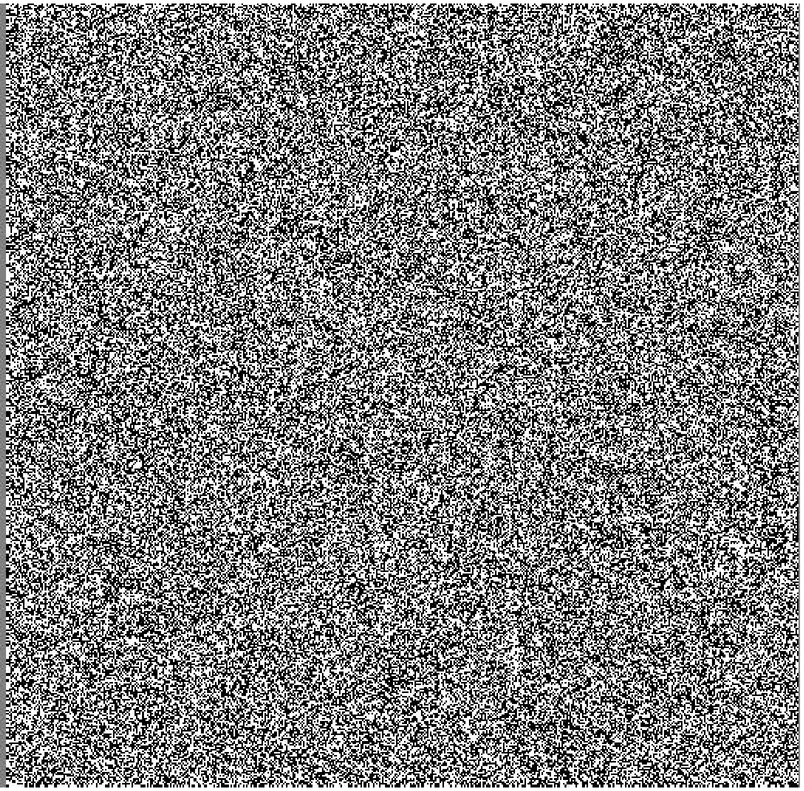
Bit Plane Representation

An image with n x m pixels each with a bit-depth of N bits can be represented by N binary bit planes of size n x m ranging from the most significant bit (MSB) to the least significant bit (LSB).

$x(k,l) = 113 = 01110001$



Example: Lena MSB and LSB Planes



Most significant bit plane | Least significant bit plane

Examples of Digital Still Images

Application	Pixels	Lines	bpp	Filesize
Facsimile	1728	1100	1	240 KB
Icons	192	128	24	72 KB
Still-Video	768	512	24	1.1 MB
Kodak Picture CD	1536	1024	24	4.5 MB
LS 633 (Kodak)	2041	1533	24	8.9 MB
LS-443 (Kodak)	2448	1632	24	11.4 MB
Kodak Photo CD	3072	2048	24	18 MB
DCS Pro 14n	4500	3000	36	61.7 MB

Examples of Digital Video Sequences

Application	Pixels	Lines	bpp	f/s	Data Rate
Video Phone 4:2:0	176	144	12	15	540 KB/s
SIF – 4:2:0	352	240	12	30	3.6 MB/s
CCIR 601 - 4:2:0	720	480	12	30	15 MB/s
CCIR 601 - 4:4:4	720	480	24	30	30 MB/s
HDTV (Interlace)	1920	1080	24	30	180 MB/s
HDTV (progressive)	1280	720	24	60	160 MB/s
Cineon	3656	2664	30	24	1 GB/s

Examples of Storage Media

Optical

Magnetic

Solid State

Examples of Storage Media

Media	Type	Capacity
Floppy Disk	Magnetic	1.4 MB
Zip Disk	Magnetic	100 & 250 MB
Jaz Disk	Magnetic	1 & 2 GB
8mm Cartridge	Magnetic	Up to 70 GB
D2C	Magnetic	25 GB
CD-ROM	Optical	650 MB
DVD-120mm	Optical	4.7 GB
DVD-80mm	Optical	1.9 GB
Compact Flash	Solid State	1-200 MB
Memory Stick	Solid State	32 & 64 MB

Examples of Transmission Media

Media	Bandwidth
Dial-Up Telephone	33.6-56 Kb/s
ISDN (Integrated Services Digital Network)	64 Kb/s
T1 (24 x ISDN)	1.544 Mb/s
Ethernet	10 Mb/s
Fast Ethernet	100 Mb/s
Ultra-Fast Ethernet	1 Gb/s
Cable Modem	1-2 Mb/s
FDDI (Fiber Distributed Data Interchange)	100 Mb/s

Examples of Transmission Media

Media	Bandwidth
Cell Phone (US)	9.6 – 14.4 Kb/s
IR (Infra-Red) Line-of-sight	115 kb/s – 16 Mb/s
Bluetooth RF	400 Kb/s
Bluetooth RF (Asymmetric)	741 Kb/s Downstream 56 Kb/s Upstream
Kodak RF	16 Mb/s
DSL (Asymmetric)	384 Kb/s – 1.5 Mb/s Downstream 128 Kb/s – 750 Kb/s Upstream

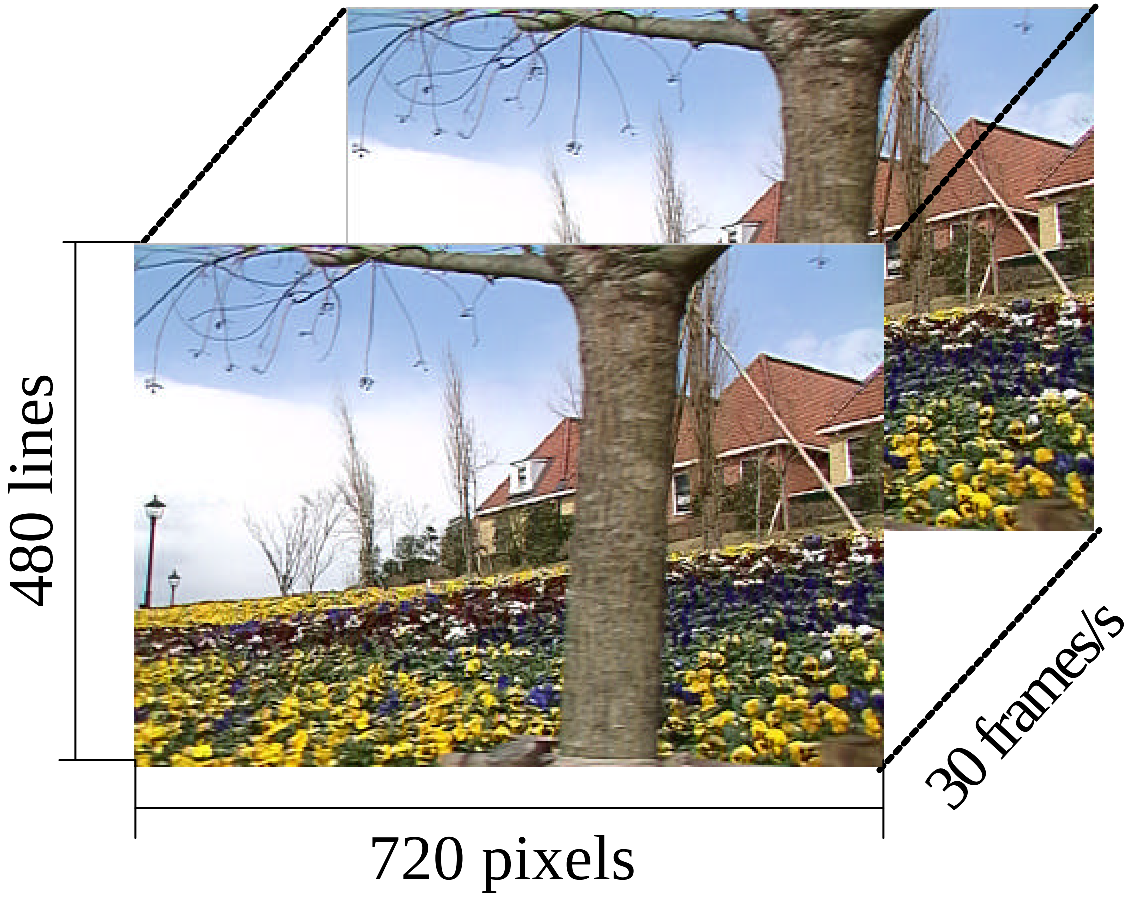
Image Transmission Example

Media	Time seconds
Cell Phone	417
28.8 Modem	139
ISDN	62
Asymmetric DSL (Downstream)	10
Bluetooth	5.5
Cable Modem (1 Mb/s)	4
IR (4 Mb/s)	1
RF (16 Mb/s)	1/4

For an uncompressed VGA 640 x 480 (4:2:0) image ~500KB

Movies on DVD

- The full resolution RGB uncompressed DVD movie has a rate of ~60 MB/sec, i.e., a DVD disc holds ~1 minute of uncompressed video!



Aerial Photograph

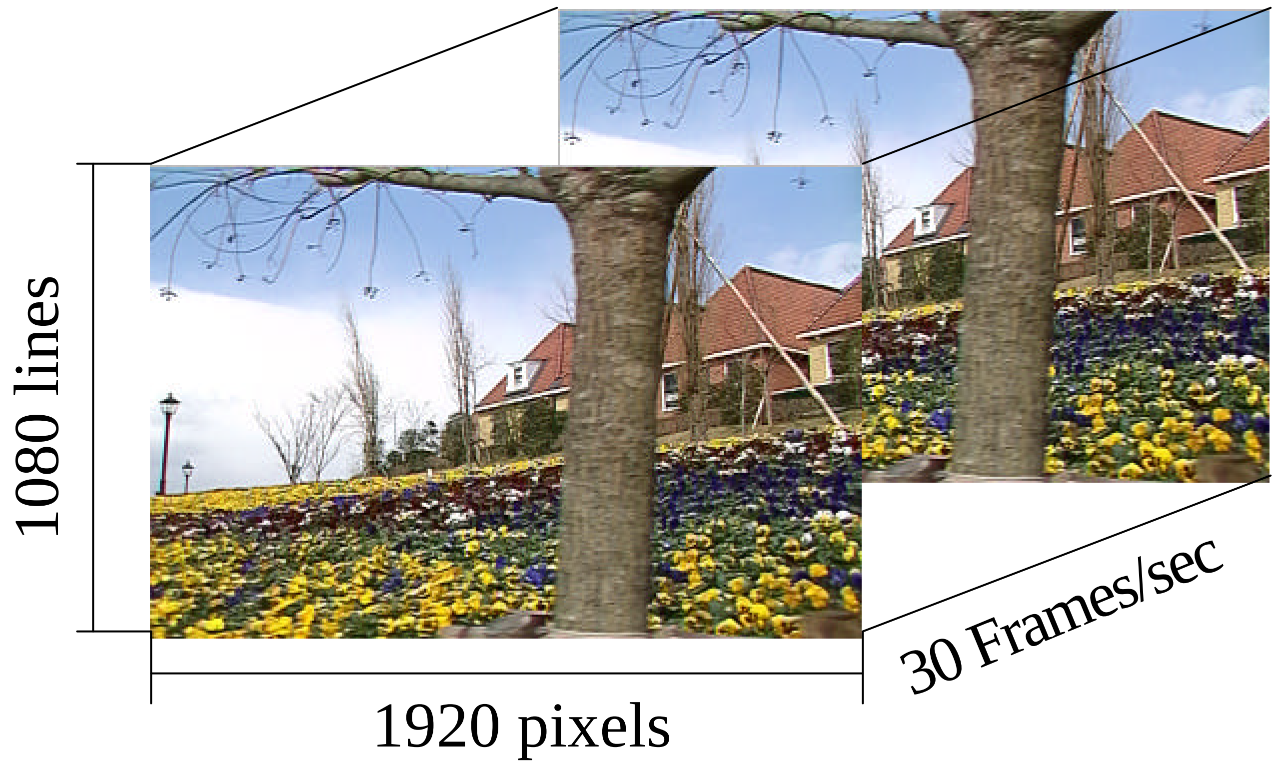
- 10” Aerial Mapping Camera at about 6” resolution
- 2.35 Square Miles of Coverage (about 1.5 Miles on a side)
- 16,000-by-16,000 pixels, 3 bands, 8 bits per band
- 730 MB uncompressed image does not fit on CD

Satellite Image

- IKONOS or QuickBird Images
- ~ 10K-by-10K, 5 bands, 10 bits per pixel

HDTV Transmission

- 1920 pixels × 1080 lines × 3 colors × 30 f/s ~ 180 MB/s
- At 18 Mb/s, a compression ratio of ~80:1 is required.



Need For Image Compression

Application	Size / Rate	CR	Method
Digital Camera	1-10 MB	>10	JPEG
Facsimile	240 KB	20	G3 & G4
CD-I (SIF 4:2:0)	3.6 MB/s	25	MPEG 1
DVD (CCIR 601 4:2:0)	15 MB/s	25	MPEG 2
HDTV	89 MB/s	70	MPEG 2
Teleconferencing	9 MB/s	50	H.261
Video Phone (QCIF)	0.54 MB/s	80	H.263

Image Compression

- Aims at finding methods for reducing the number of bits needed to represent a digital image without compromising the required image quality for a given application.
- Image compression is used to:
  - Reduce memory for image storage.
  - Reduce the bandwidth/time required for image transmission.
  - Increase effective data transfer rate.

Why images can be compressed

What is digital image compression

- Image compression is the art/science of finding efficient representations for digital images in order to:
  - Reduce the memory required for their storage,
  - Reduce the effective data access time when reading from storage devices,
  - Reduce the bandwidth and/or the time required for their transfer across communication channels.

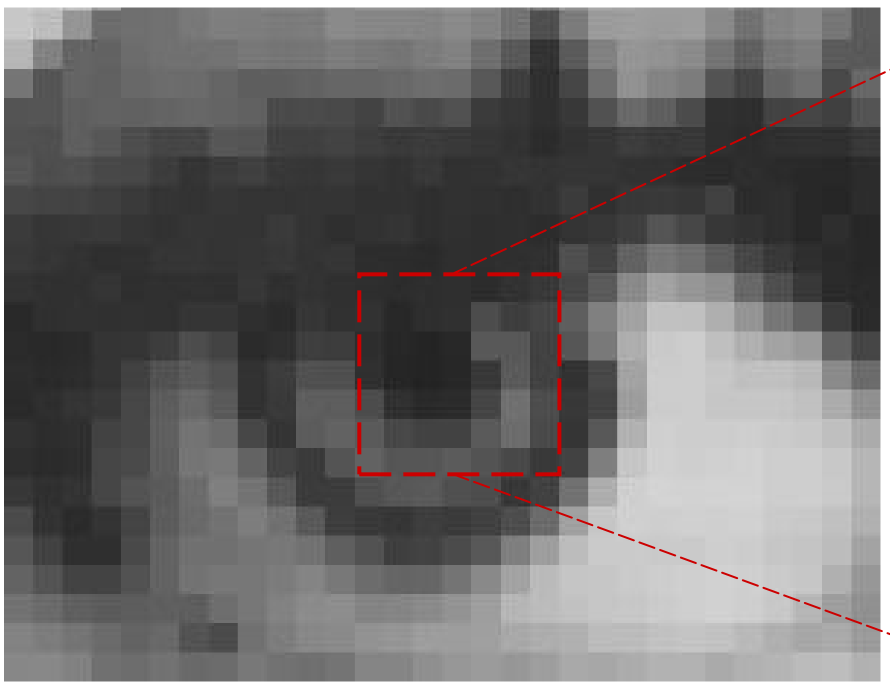
- The goal is to achieve the desired bit rate without compromising the image quality required for a given application.

Why can images be compressed?

- Image compression can be achieved because image data are often highly redundant and/or irrelevant. The redundancy or irrelevancy exist in **spatial**, **spectral**, and **temporal** forms.
  - **Redundancy** relates to the statistical properties of an image (e.g., pixel-to-pixel correlation, spectral (RGB) correlation, frame-to-frame similarity, etc.) and is a function of resolution, bit-depth, image noise, and image detail.
  - **Irrelevancy** relates to an observer viewing an image (HVS spatial and temporal CSF, visual masking, etc.) and is a function of image resolution, noise, detail, and viewing conditions.

Spatial Redundancy And Irrelevancy

- What is the value of the missing pixel? **(39)**
- How critical is its exact reproduction?



46	44	46	47	44	50	56
50	40	46	48	76	62	69
45	39	36	40	88	87	65
48	38	36	?	56	90	65
76	48	41	43	69	112	77
90	72	66	66	90	108	74
98	86	86	91	83	72	57

Example: Original and Difference Images

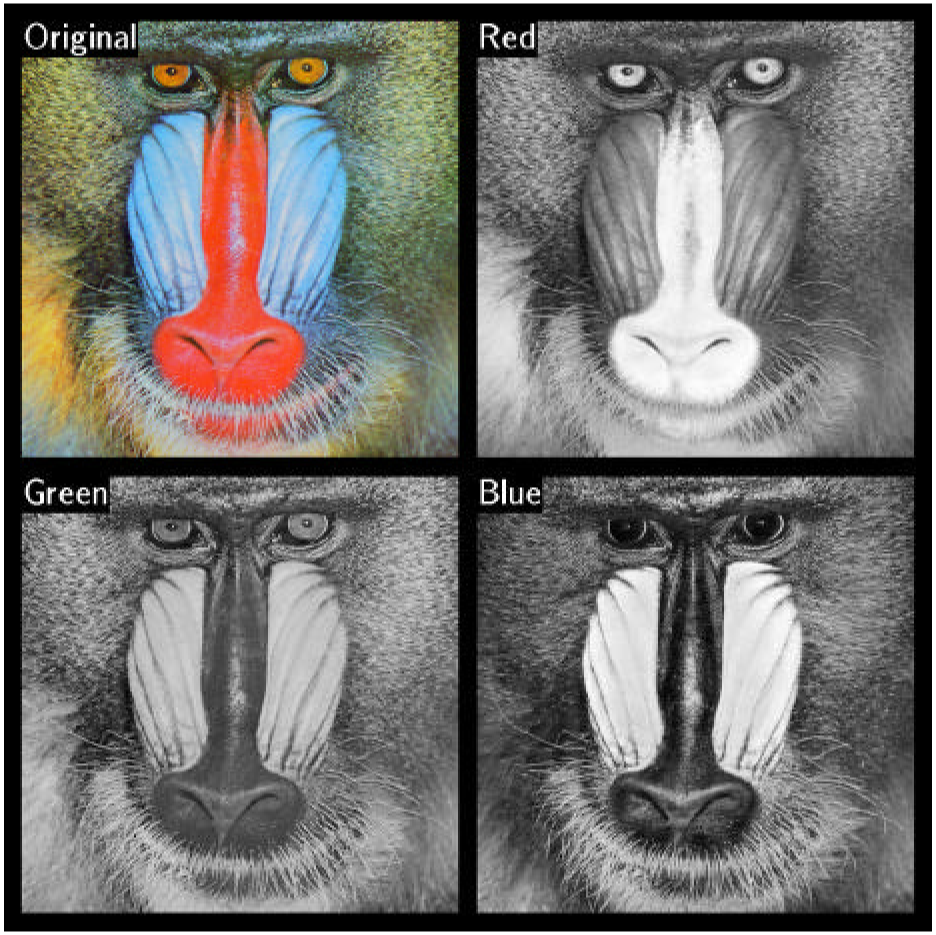
139 139	144 +5	149 +5	153 +4	155 +2	155 0	155 0	155 0
144 +5	151 +7	153 +2	156 +3	159 +3	156 -3	156 0	156 0
150 +6	155 +5	160 +5	163 +3	158 -5	156 -2	155 0	155 0
159 +9	161 +2	162 +1	160 -2	160 0	159 -1	159 0	159 0
159 0	160 +1	161 +1	162 +1	162 0	155 -7	155 0	155 0
161 +2	161 0	161 0	161 0	160 -1	157 -3	157 0	157 0
162 +1	162 0	161 -1	163 +2	162 -1	157 -5	157 0	157 0
162 0	162 0	161 -1	161 0	163 +2	158 -5	158 0	158 0

Spatial Redundancy

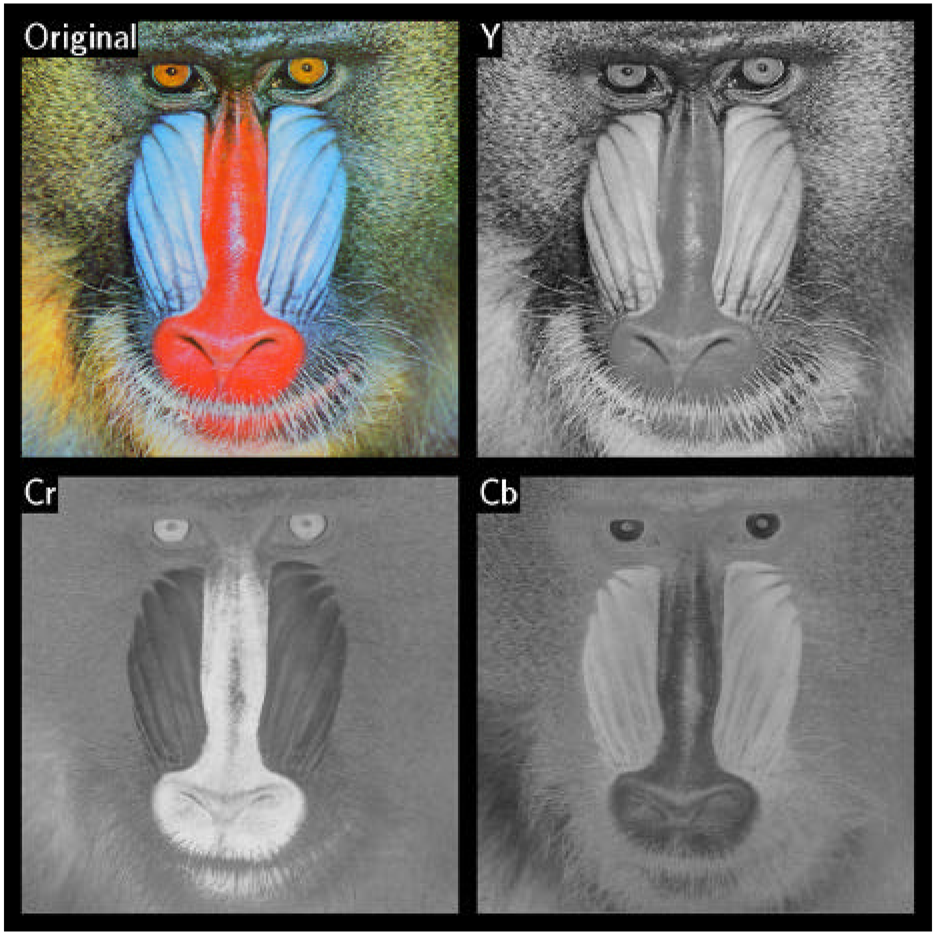
- The difference between two adjacent pixels has a very skewed distribution centered around zero.

Spectral Redundancy

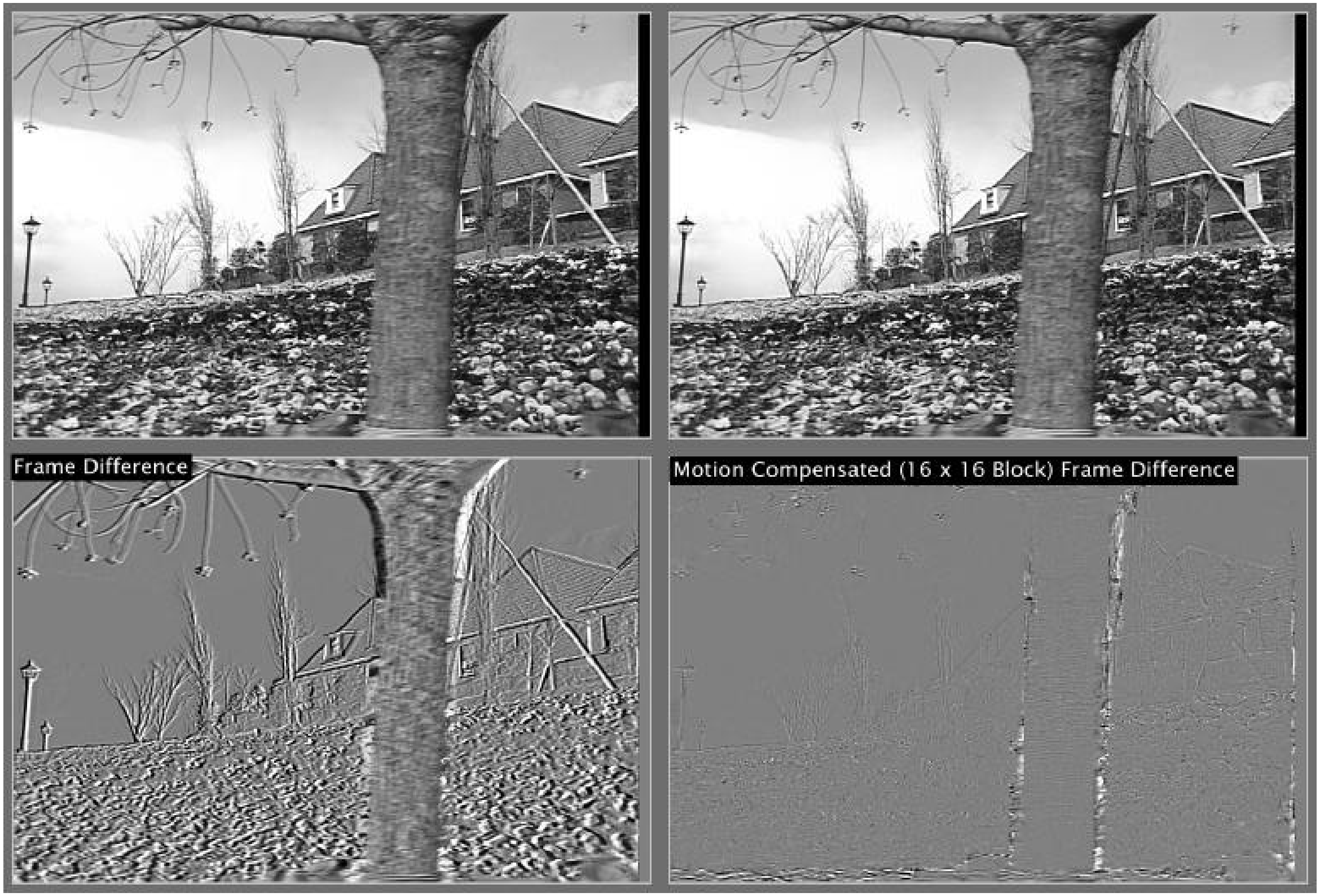
Original RGB



Transformed YCrCb



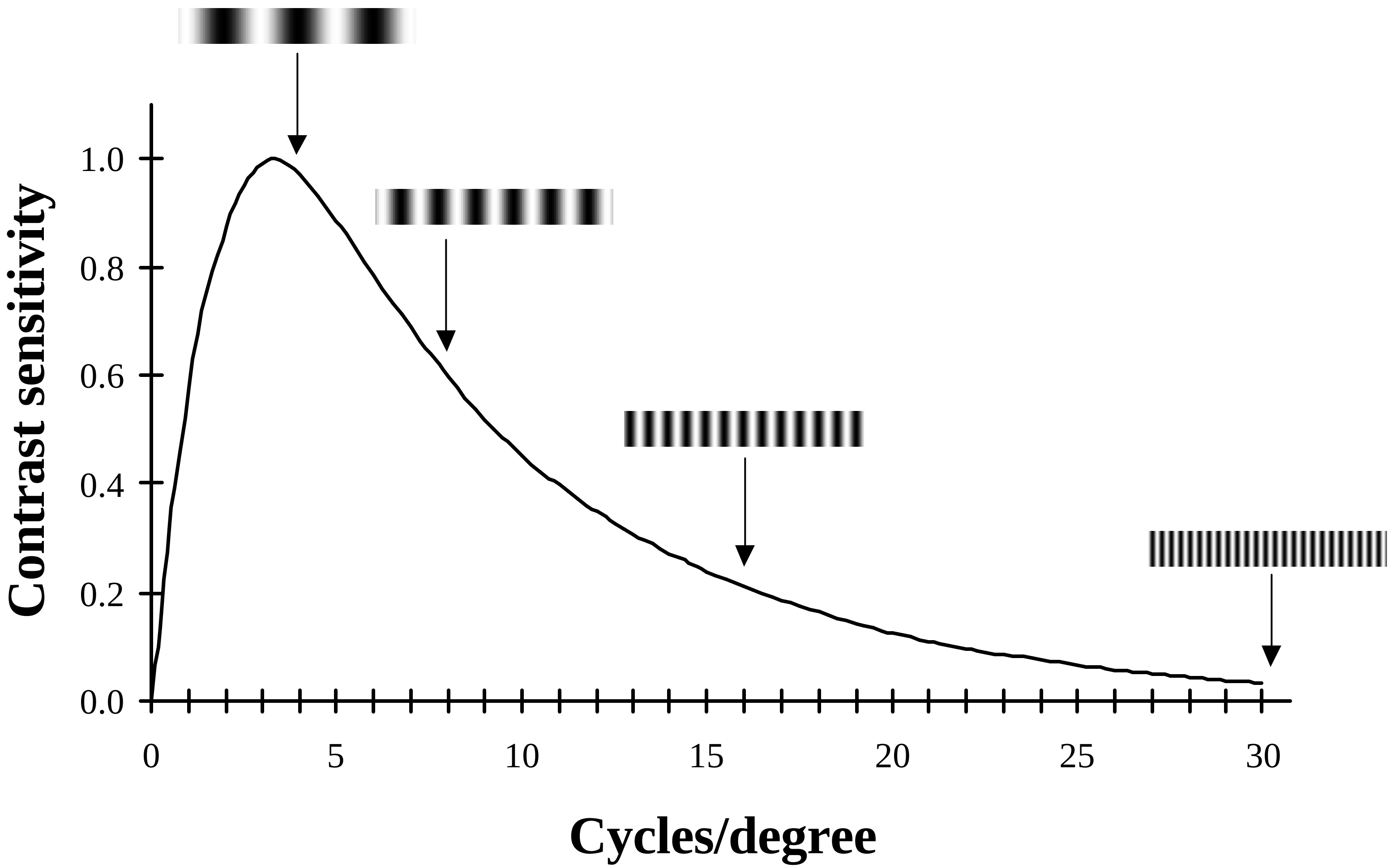
Temporal Redundancy

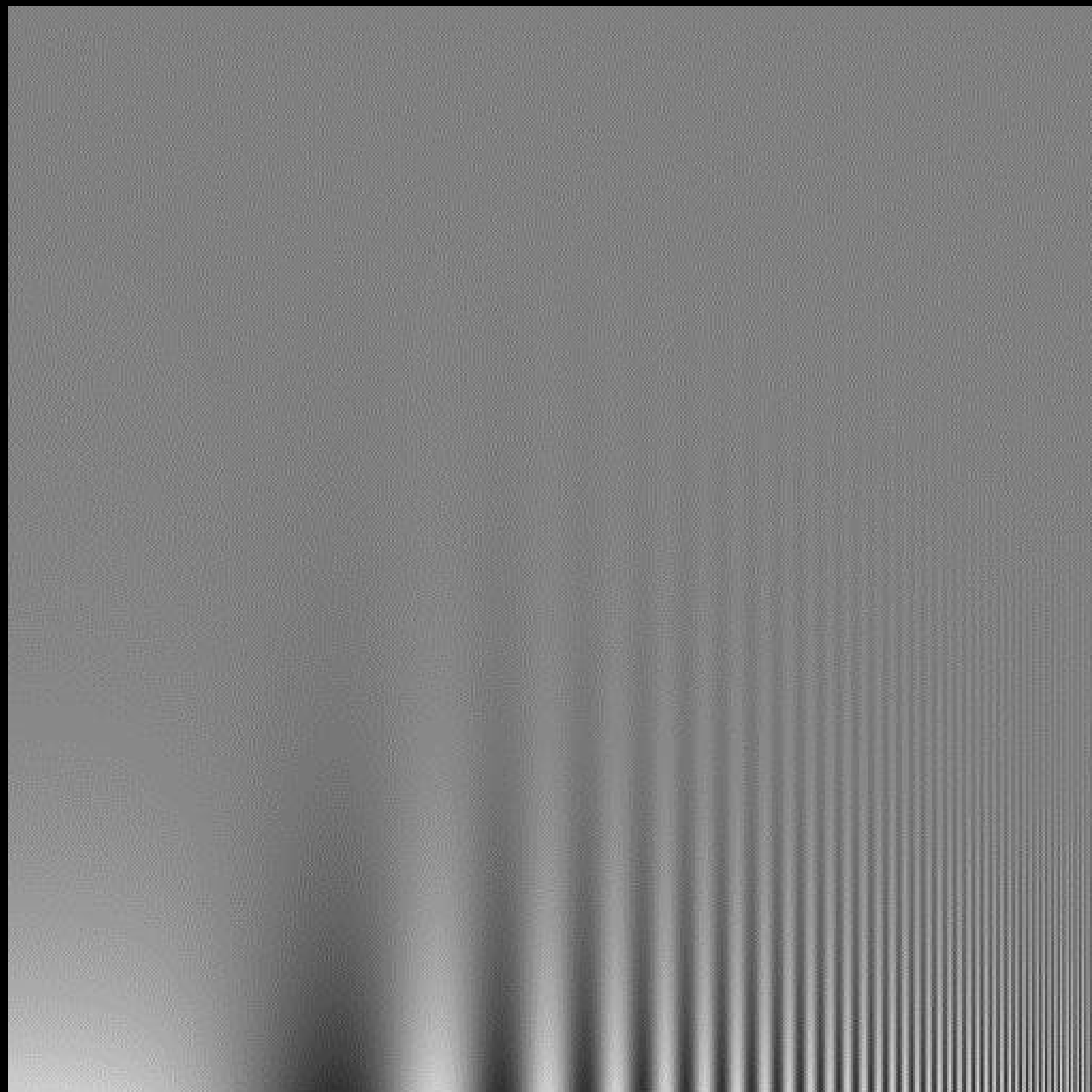


The Human Visual System CSF

- The frequency-dependent behavior of the human visual system (**HVS**) can be characterized by its response to harmonic (sinusoidal) functions.
- For each sinusoid with a given frequency, the amount of contrast needed to elicit a criterion level of response from a neuron is called the **contrast threshold**.
- The inverse of the contrast threshold is called the **contrast sensitivity** and when plotted as a function of frequency is referred to as the **contrast sensitivity function (CSF)**.
- The luminance CSF peaks at around 5 cycles/degree, and rapidly drops off to almost zero at 50 cycles/degree. The chrominance CSF drops even faster.

Example of Luminance CSF

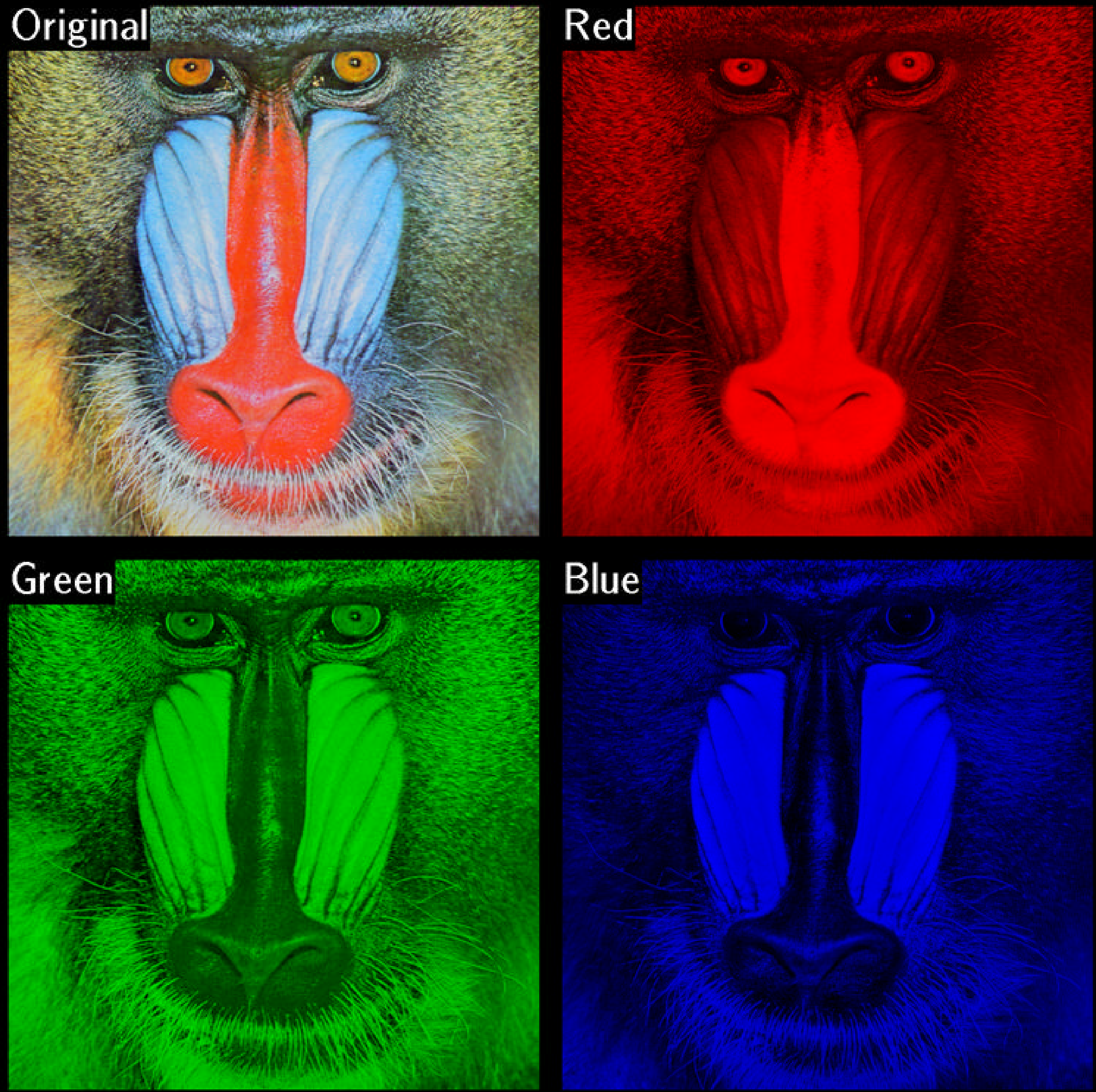


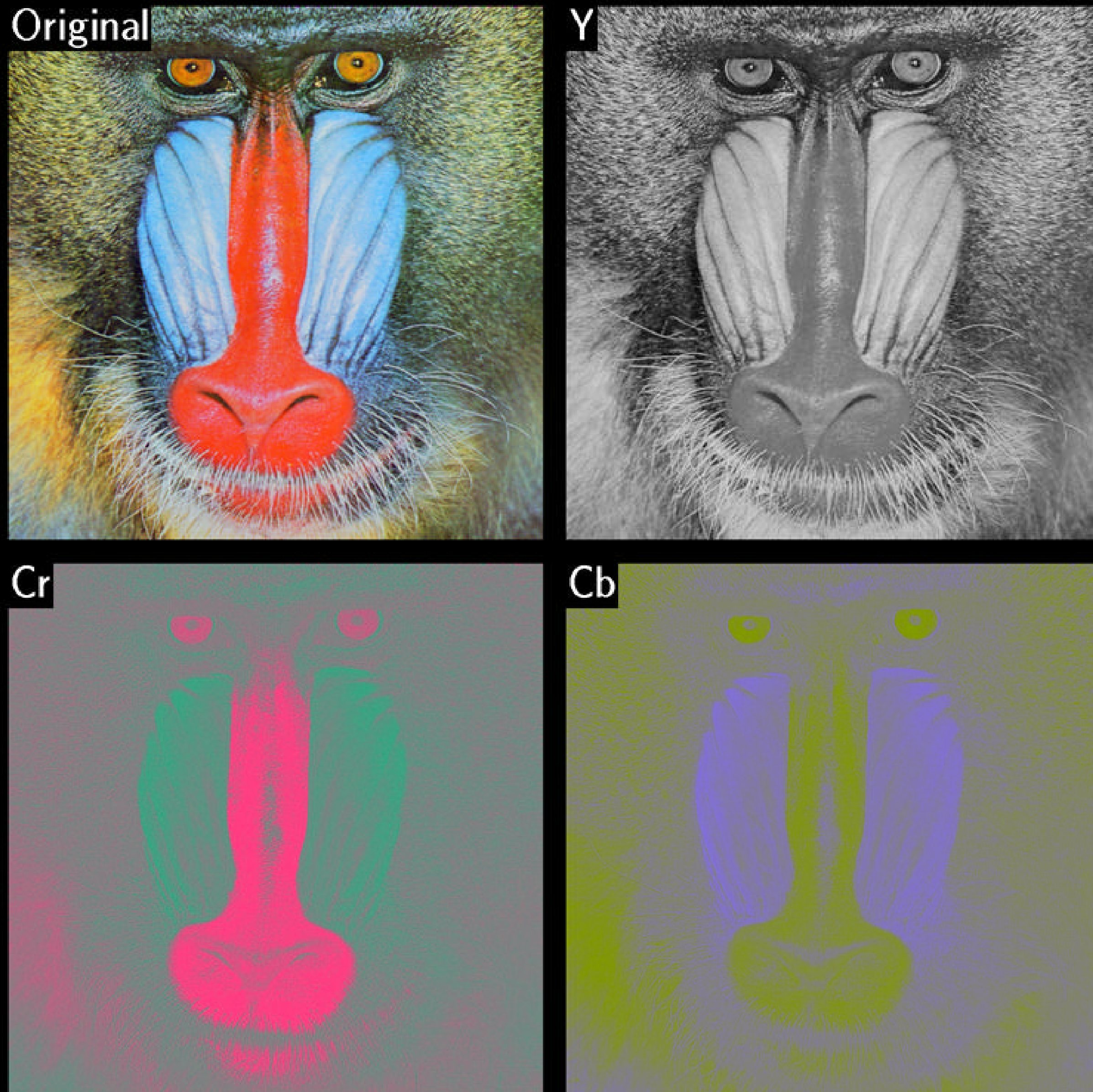



Original (512x512)


Original + Distortion


Distortion

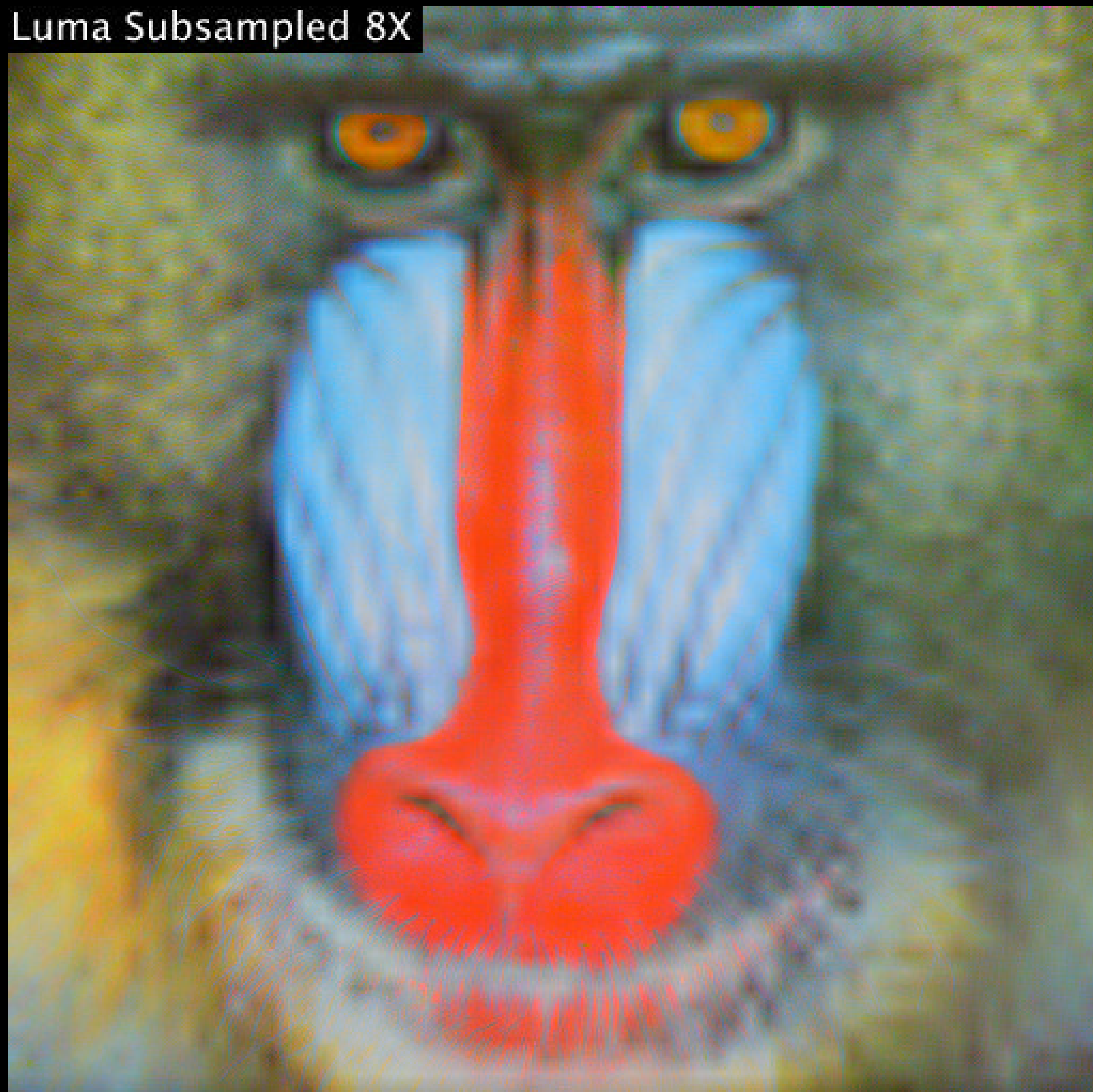

Original
Red
Green
Blue


Original
Y
Cr
Cb

Kodak


Luma Subsampled 2X


Luma Subsampled 4X


Luma Subsampled 8X


Chroma Subsampled 2X


Chroma Subsampled 4X

Chroma Subsampled 8X

Image Compression Basics

Basic Strategy in Image Compression

- Ideally, an image compression technique removes redundant and/or irrelevant information, and efficiently encodes what remains.
- Practically, it is often necessary to throw away both non-redundant information and relevant information to achieve the required compression.
- In either case, the trick is finding methods that allow important information to be efficiently extracted and represented.

Some Factors Affecting Achievable Compression

- Sample parameters (spatial resolution, bit depth).
- Sensor characteristics (noise, spectral response).
- Scene content, including noise.
- Image size and viewing distance.
- Display characteristics (noise, light level, non-linearities)
- Post Processing (Sharpening, Dynamic Range Adjustment (DRA), Tone Transfer Curve (TTC))
- Pre-Processing (image formation, registration)
- Observer (IA, machine)
- Required task

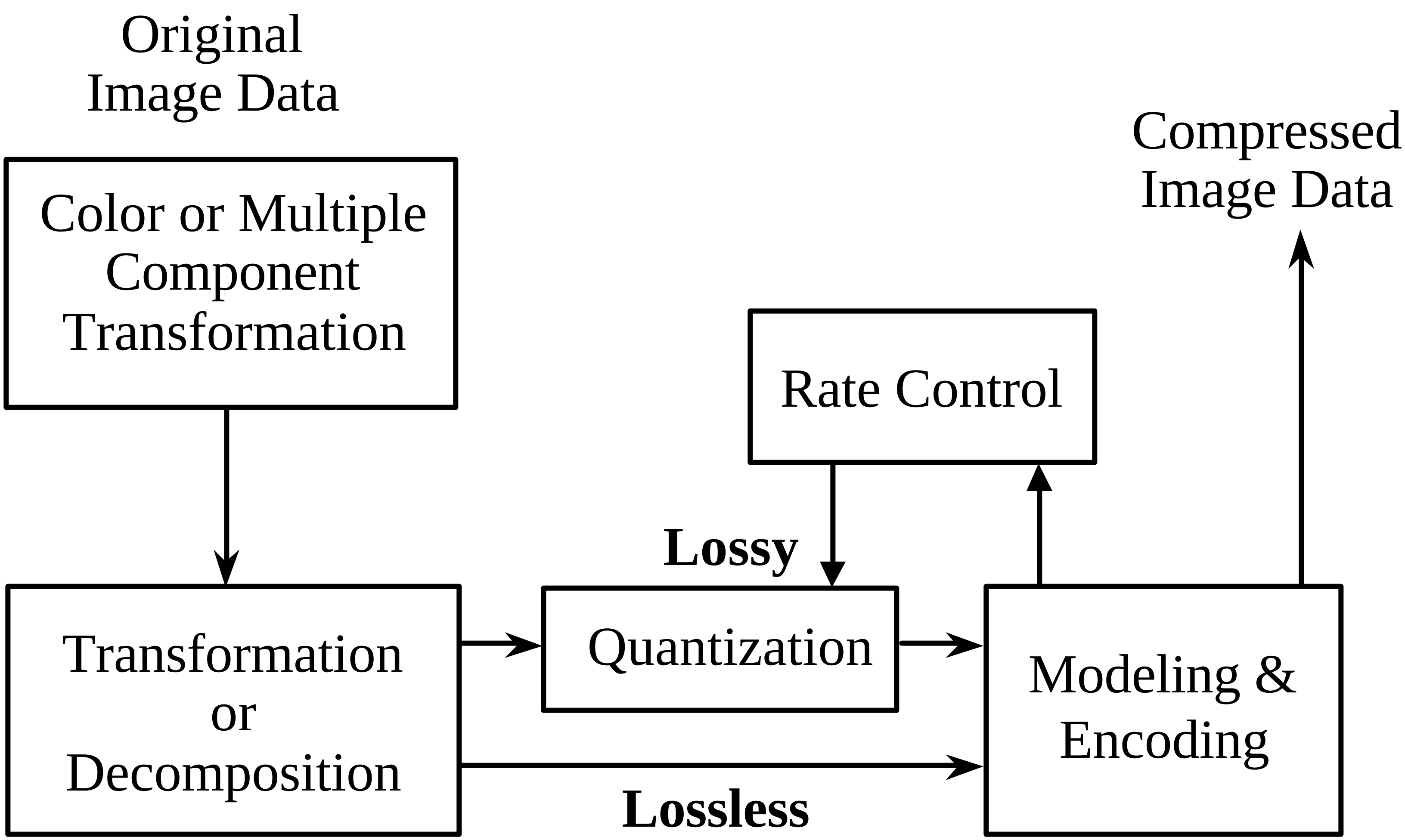
Lossless (Reversible) Compression

- The image after compression and decompression is identical to the original.
- Only the statistical redundancy is exploited to achieve compression.
- Data compression techniques such as LZW or LZ77 are used in GIF, PNG, and TIFF file formats and the Unix “Compress” command.
- Image compression techniques such as lossless JPEG or JPEG-LS perform slightly better.
- Compression ratios are typically ~2:1 for natural imagery but can be much larger for document images.

Lossy (Irreversible) Compression

- The reconstructed image contains degradations with respect to the original image.
- Both the statistical redundancy and the perceptual irrelevancy of image data are exploited.
- Much higher compression ratios compared to lossless.
- Image quality can be traded for compression ratio.
- The term **visually lossless** is often used to characterize lossy compression schemes that result in no visible degradation under a set of designated viewing conditions..

Compression Framework



Transformation

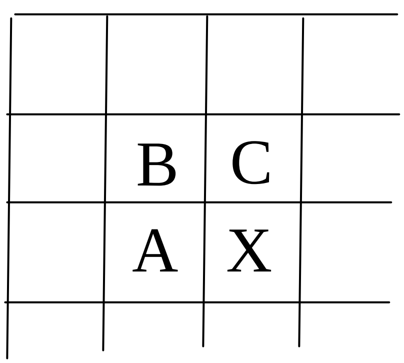
Decomposition or Transformation

- A reversible process (or near-reversible, due to finite precision arithmetic) that reduces redundancy and/or provides an image representation that is more amenable to the efficient extraction and coding of relevant information.

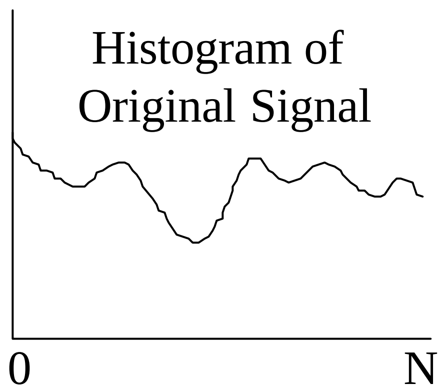
- Examples
  - Block-based linear transformations, e.g. Discrete Cosine Transform (DCT)
  - Wavelet decompositions.
  - Prediction/residual formation, e.g. Differential Pulse Code Modulation (DPCM)
  - Color space transformations, e.g. RGB to YCrCb.
  - Model prediction/residual formation, e.g. Fractals

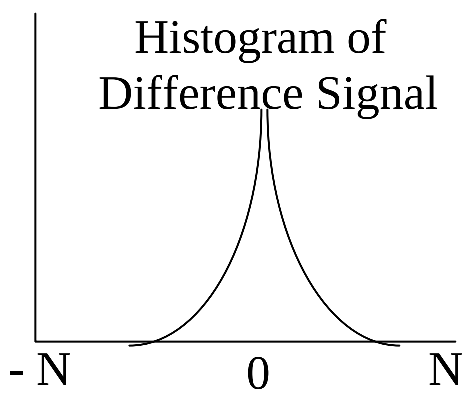
DPCM

- Lossless JPEG and 4.3 DPCM are based on differential pulse code modulation (DPCM).
  - In DPCM, a combination of previously encoded pixels (A, B, C) is used as a prediction (χ) for the current pixel (X).
  - The difference between the actual value and the prediction (χ – X) is encoded using Huffman coding.
    - The quantized difference is encoded in lossy DPCM
  - Properties
    - Low complexity
    - High quality (limited compression)
    - Low memory requirements

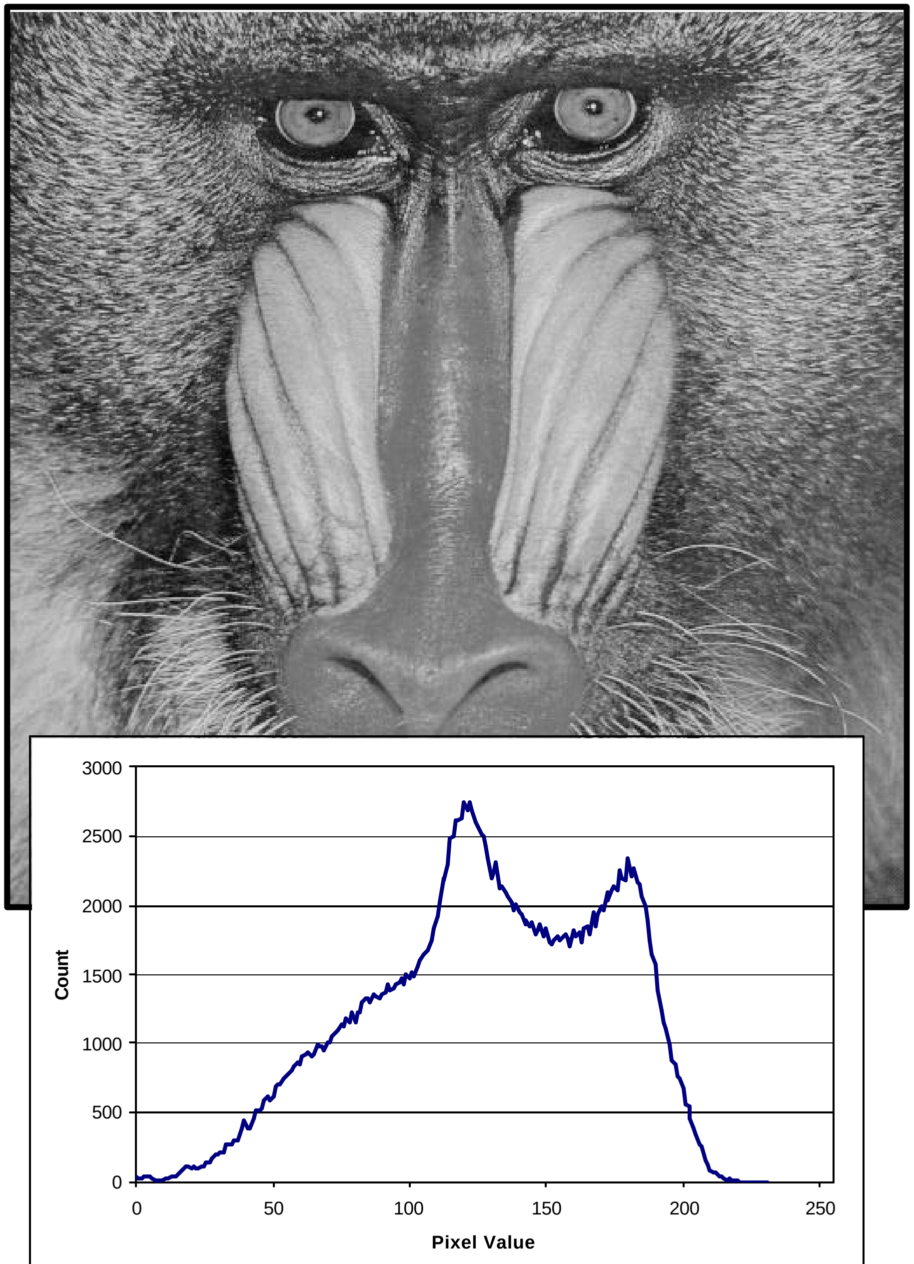


χ = A
χ = (A +C)/2
χ = (A + C - B)

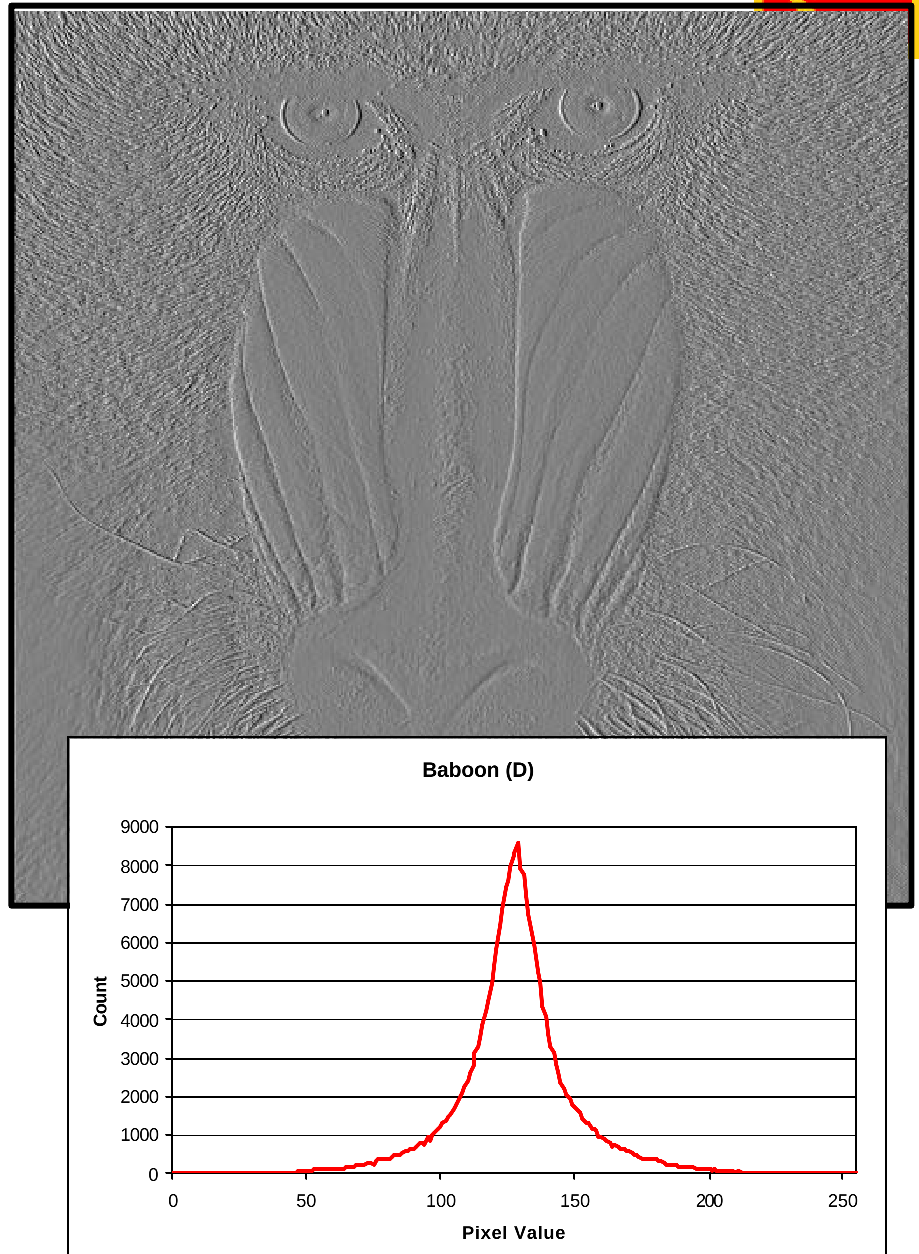




Original



DPCM output



Example Block From Lena Image

Following is an 8x8 block of the Lena image where each pixel value has been level-shifted by subtracting a value of 128.

$x(k,l) =$

8	14	23	37	52	68	73	82
6	14	24	37	46	67	74	81
3	11	28	35	48	62	72	82
4	13	22	28	44	61	69	86
5	11	18	30	40	59	72	86
5	9	16	29	39	58	74	83
-1	8	16	31	38	59	75	80
2	11	18	30	37	57	69	82

2-Dimensional 8 x 8 DCT Basis Functions

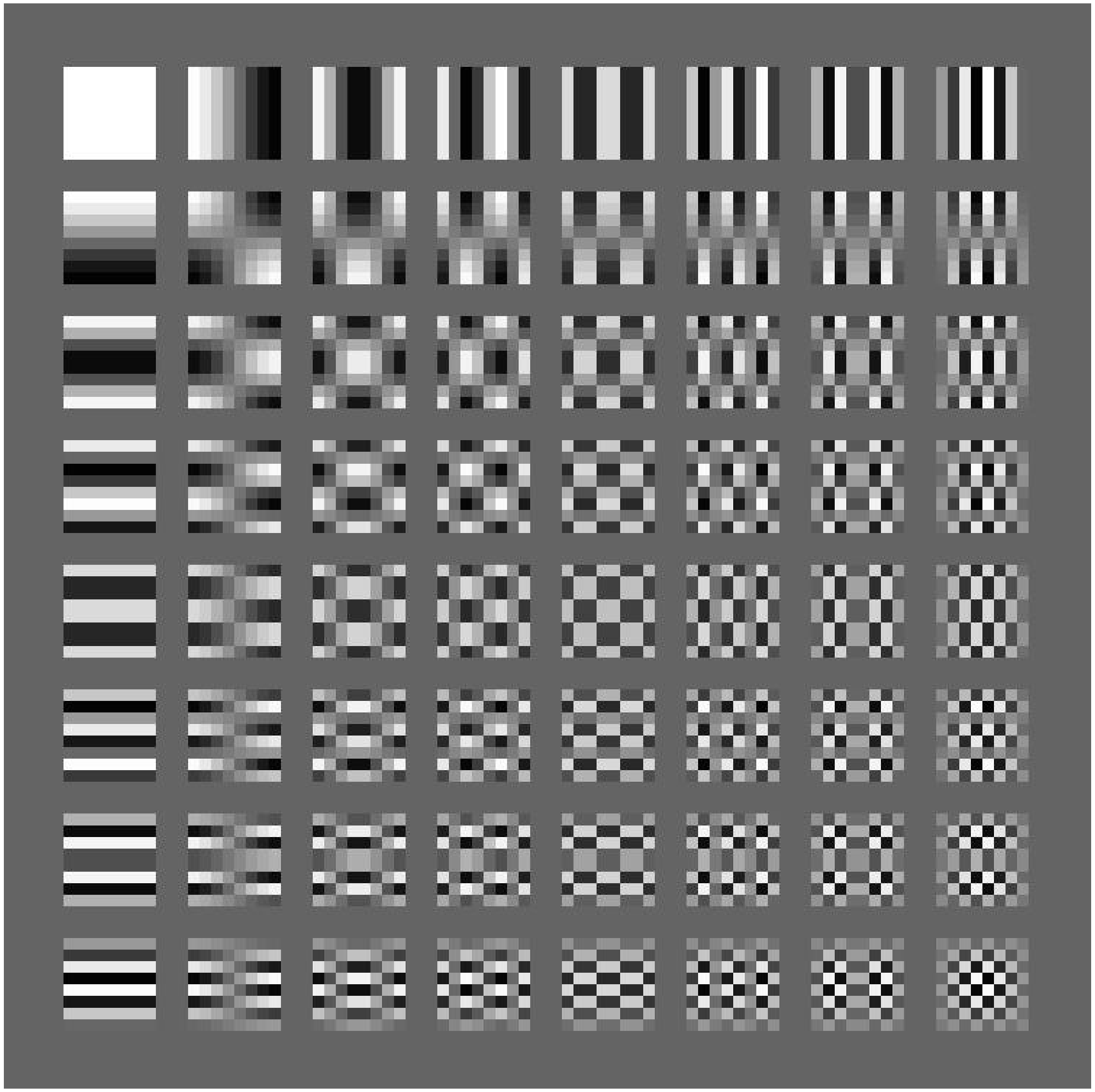
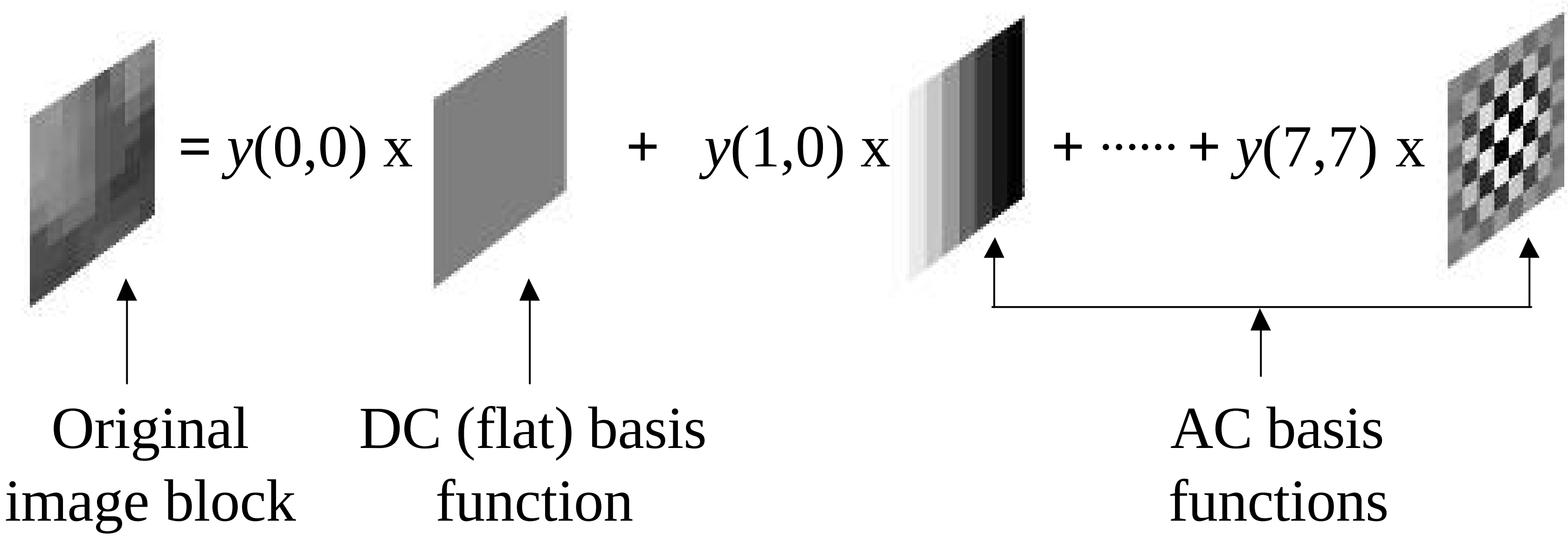


Image Representation with DCT

- DCT coefficients can be viewed as weighting functions that, when applied to the 64 cosine basis functions of various spatial frequencies (8 x 8 templates), will reconstruct the original block.



DCT of 8 x 8 Image Block

The 8 x 8 DCT of the block preserves the block's energy (sum of the squared amplitudes), but it packs the block energy into a small number of DCT coefficients by removing the pixel redundancy or correlation.

DC Value → 327.5

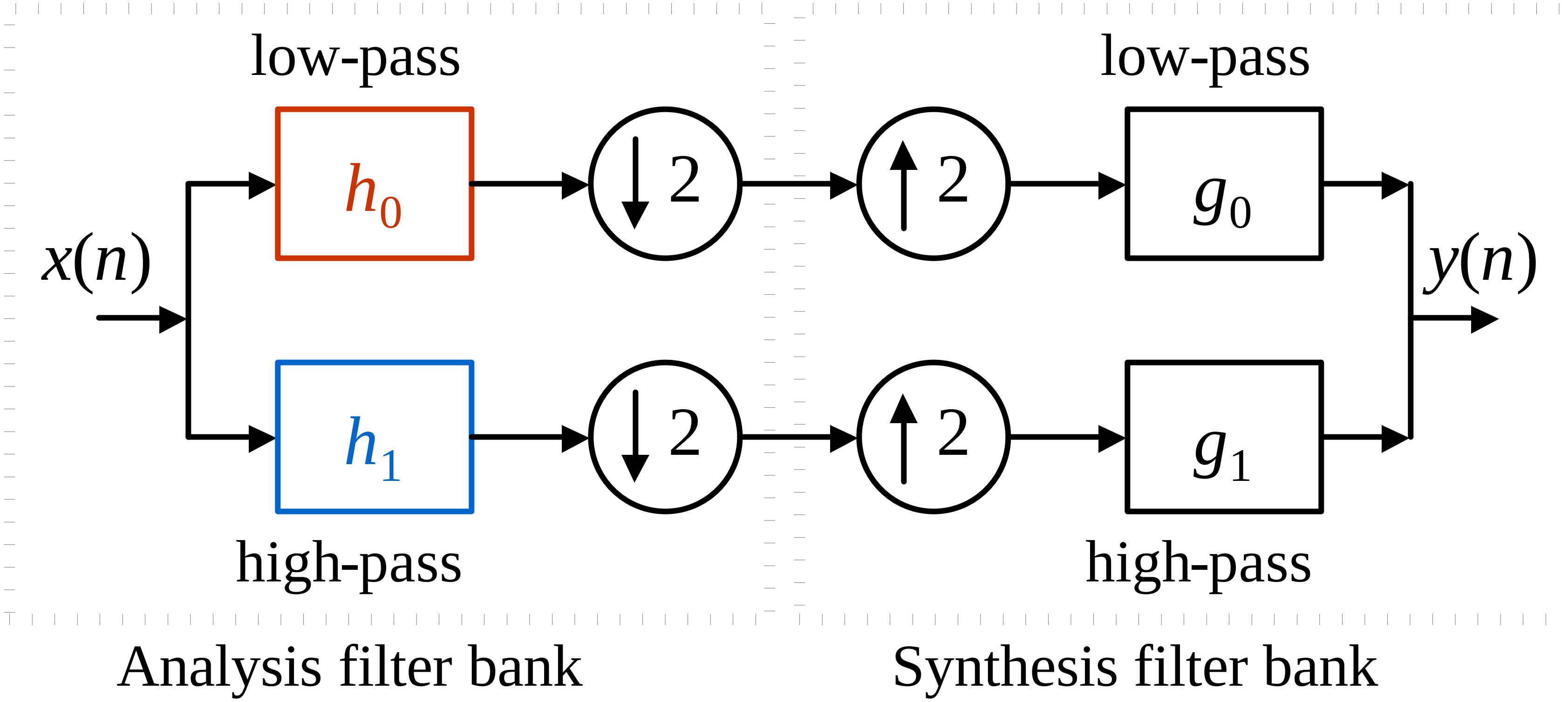
$y(u,v) =$

327.5	-215.8	16.1	-10.7	-3.7	-1.5	4.2	-6.7
18.1	3.4	-9.9	3.7	0.5	-3.2	3.5	2.2
2.5	1.3	-5.4	2.8	-1.0	2.3	-1.6	-2.6
0.6	-2.5	3.0	5.0	1.8	2.2	-2.6	-1.4
0.3	1.6	3.4	0.0	2.5	-5.1	1.6	-0.7
-0.6	-1.8	-2.4	0.5	-0.4	-1.6	-0.1	2.1
0.9	1.6	-0.6	-0.7	2.1	-0.5	0.9	2.8
0.6	-1.0	-2.9	-1.4	0.2	1.9	-0.6	0.7

1-D Discrete Wavelet Transform (DWT)

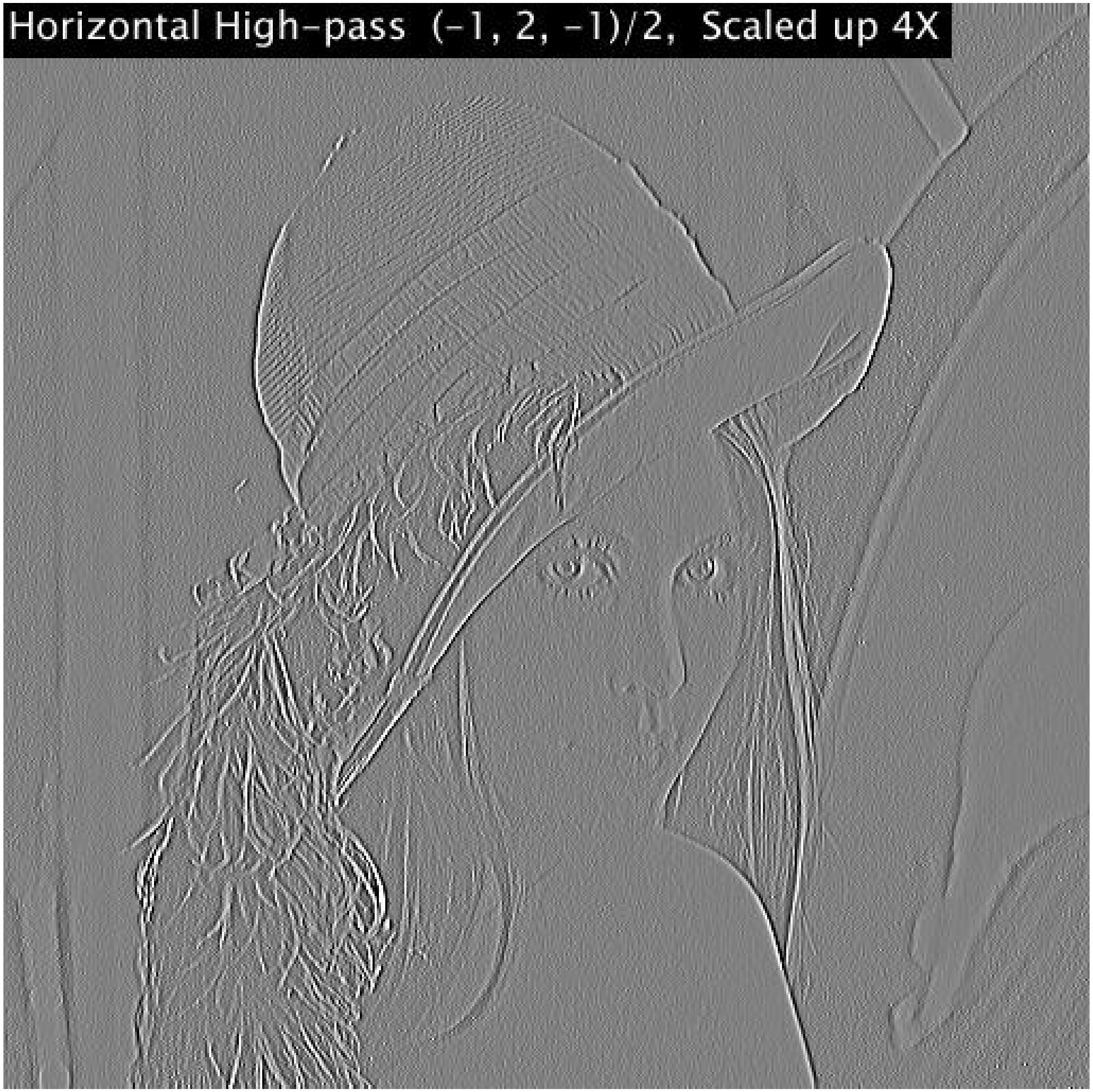
- The **forward discrete wavelet transform** (DWT) decomposes a one-dimensional (1-D) sequence (e.g., line of an image) into two sequences (called **subbands**), each with half the number of samples, according to the following procedure:
  - The 1-D sequence is separately **low-pass** and **high-pass** filtered.
  - The filtered signals are downsampled by a factor of two to form the low-pass and high-pass subbands.
  - The two filters are called the **analysis filter bank**.

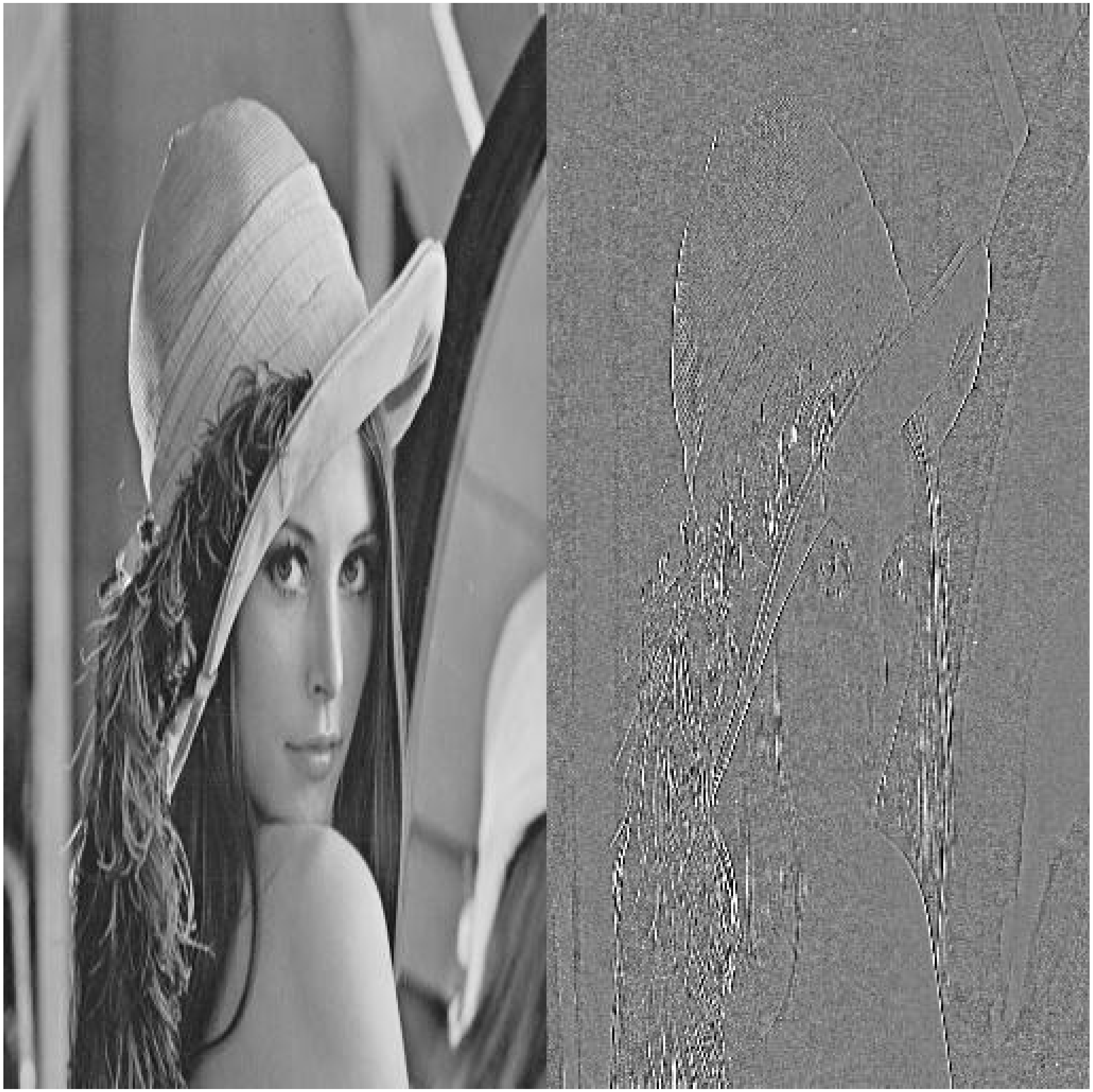
The 1-D Two-Band DWT



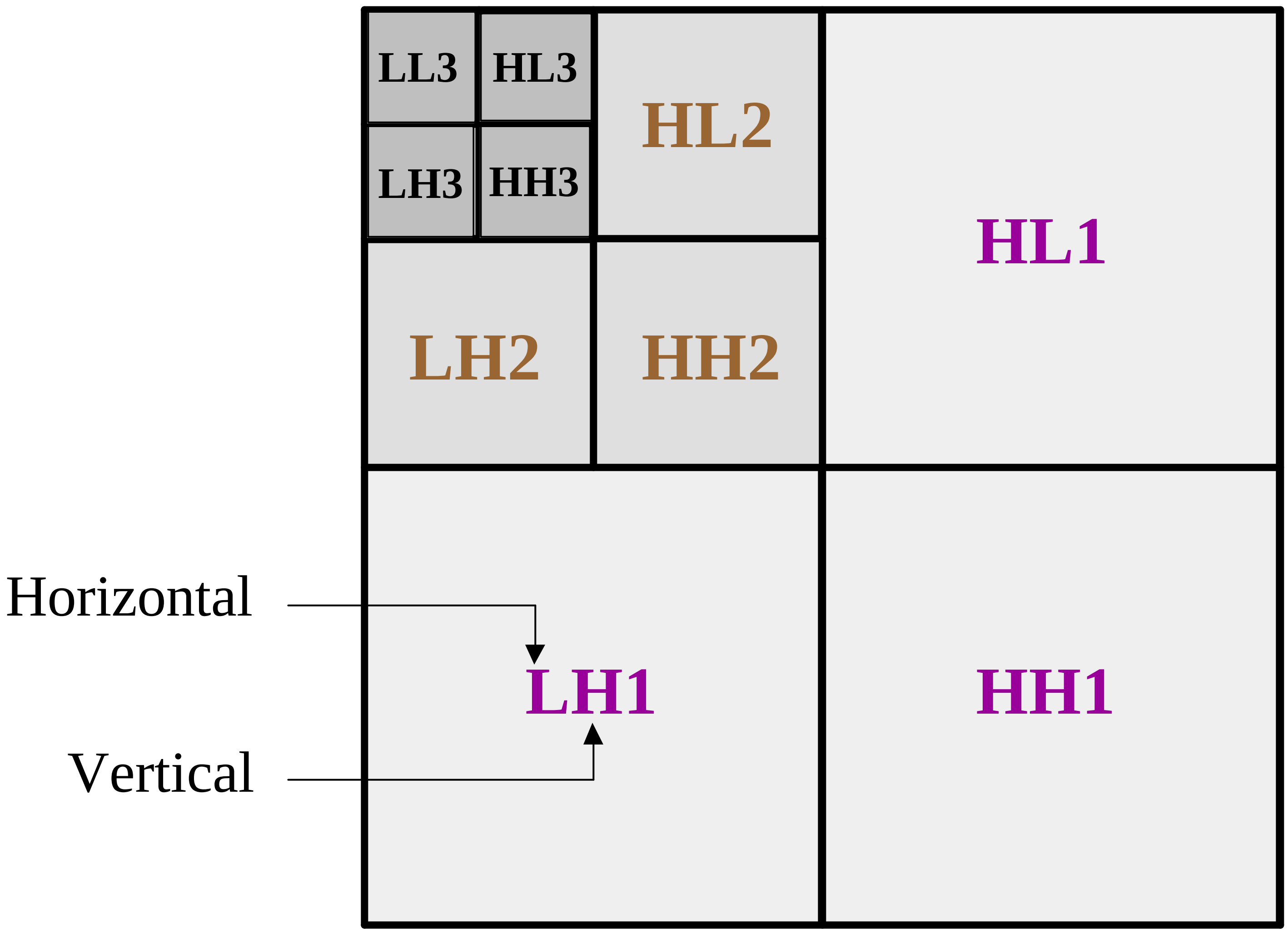
Ideally, it is desired to choose the analysis filter banks (h_0 and h_1), and the synthesis filter banks (g_0 and g_1), in such a way so as to make the overall distortion zero, i.e., $x(n) = y(n)$. This is called the **perfect reconstruction** property.


Horizontal Low-pass (-1, 2, 6, 2, -1)/8


Horizontal High-pass (-1, 2, -1)/2, Scaled up 4X



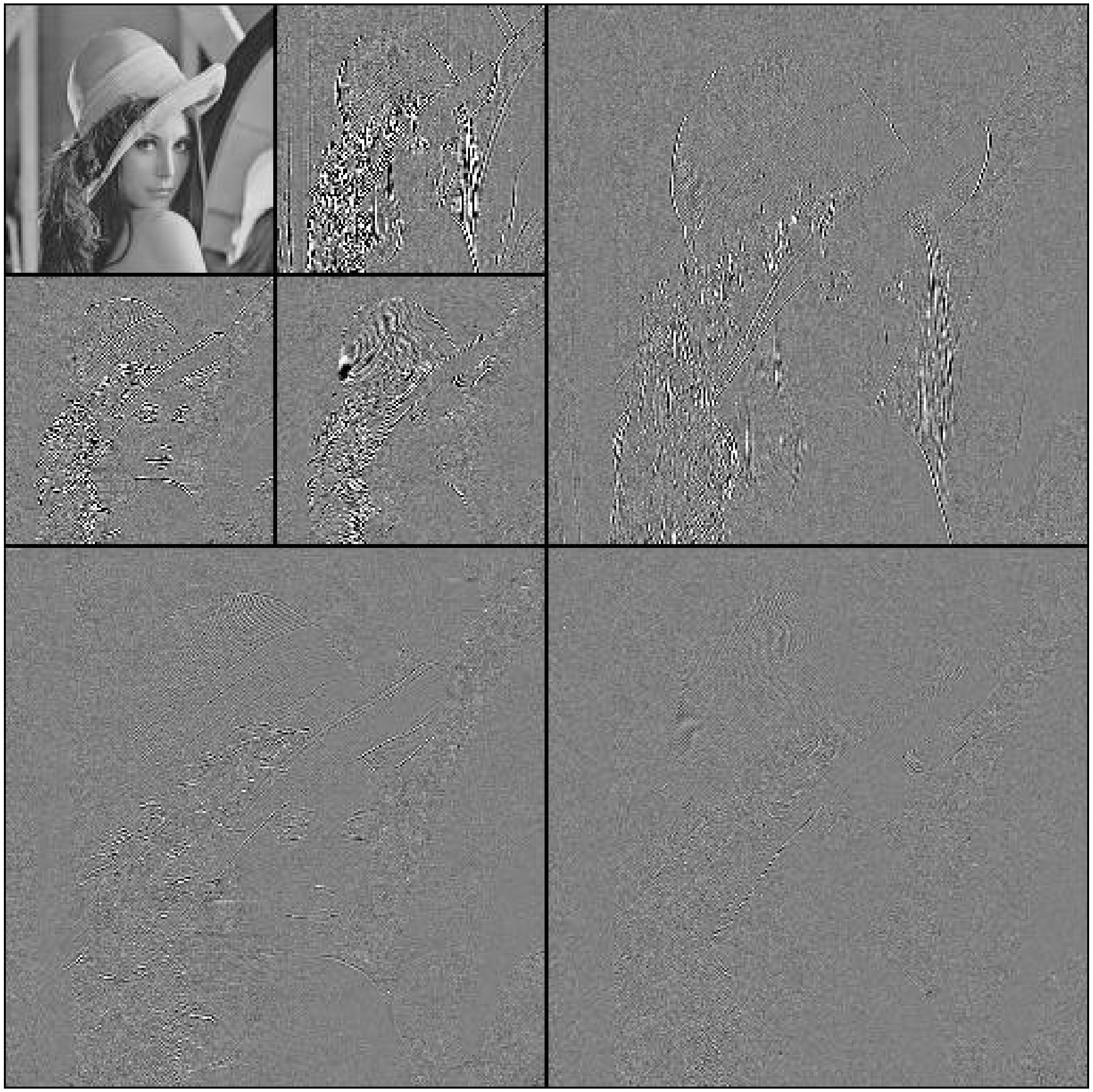
2-D Wavelet Decomposition



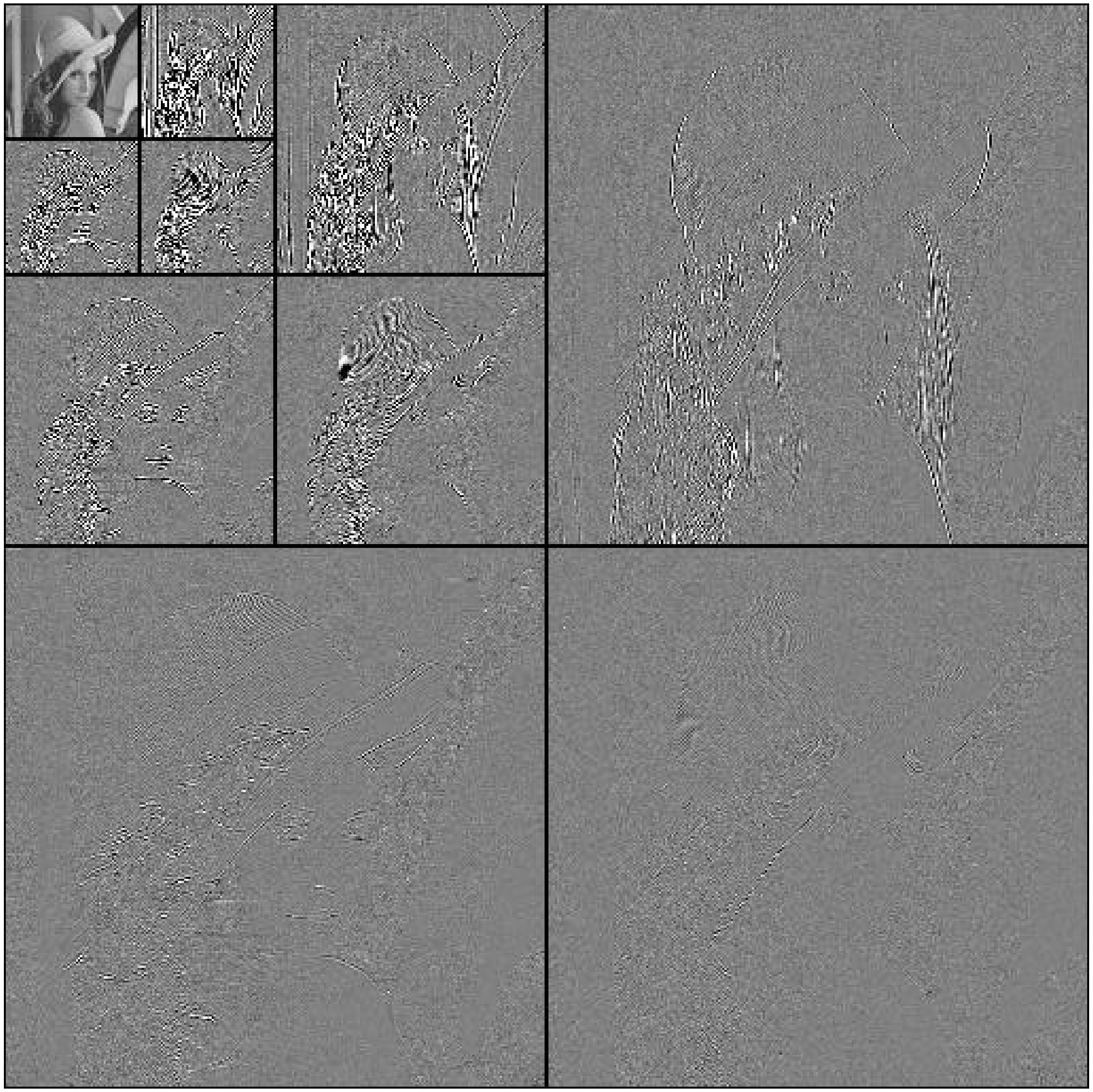
Original Lena Image

1-Level, 2-D Wavelet Decomposition of Lena

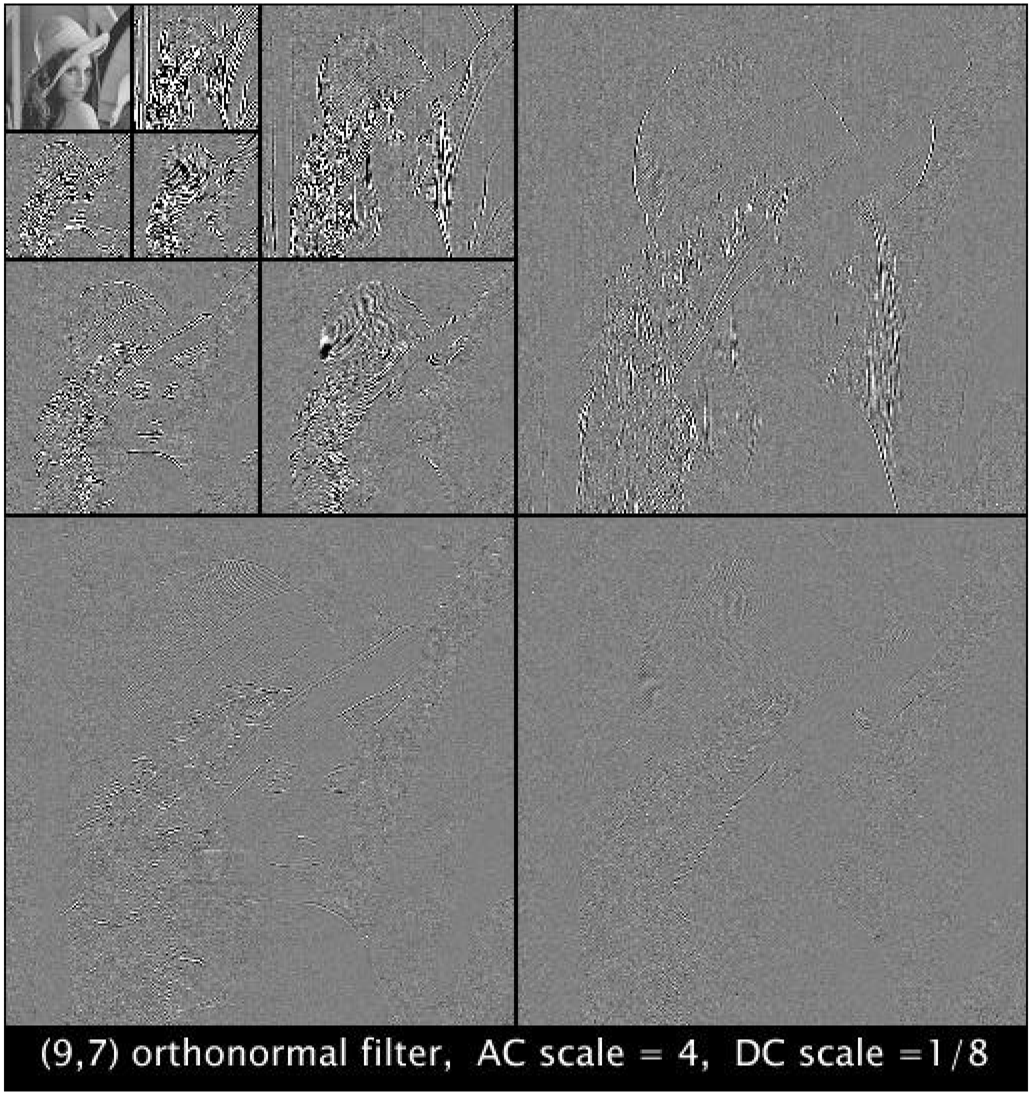
2-Level, 2-D Wavelet Decomposition of Lena



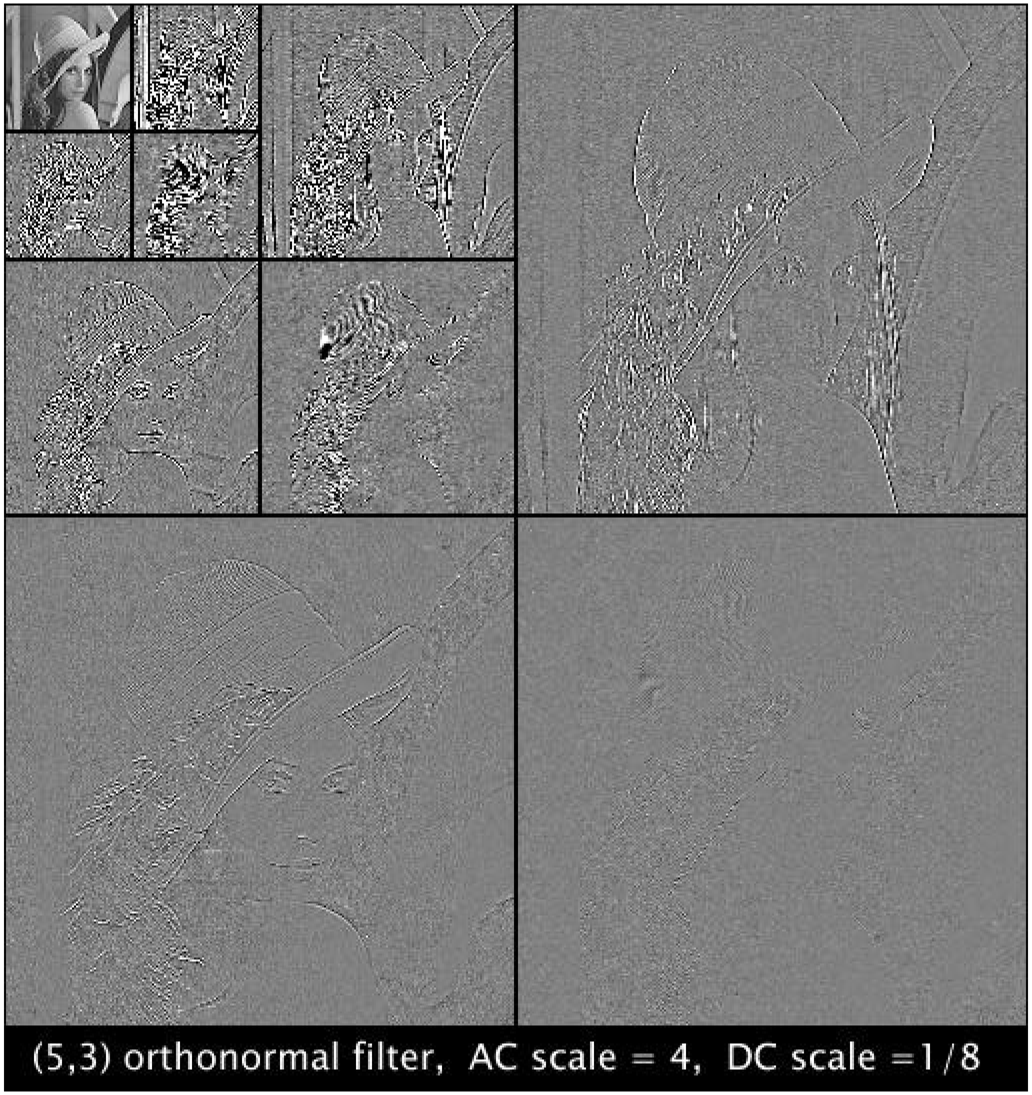
3-Level, 2-D Wavelet Decomposition of Lena



3-Level, 2-D DWT with (9,7) Filter



3-Level, 2-D DWT with (5,3) Filter

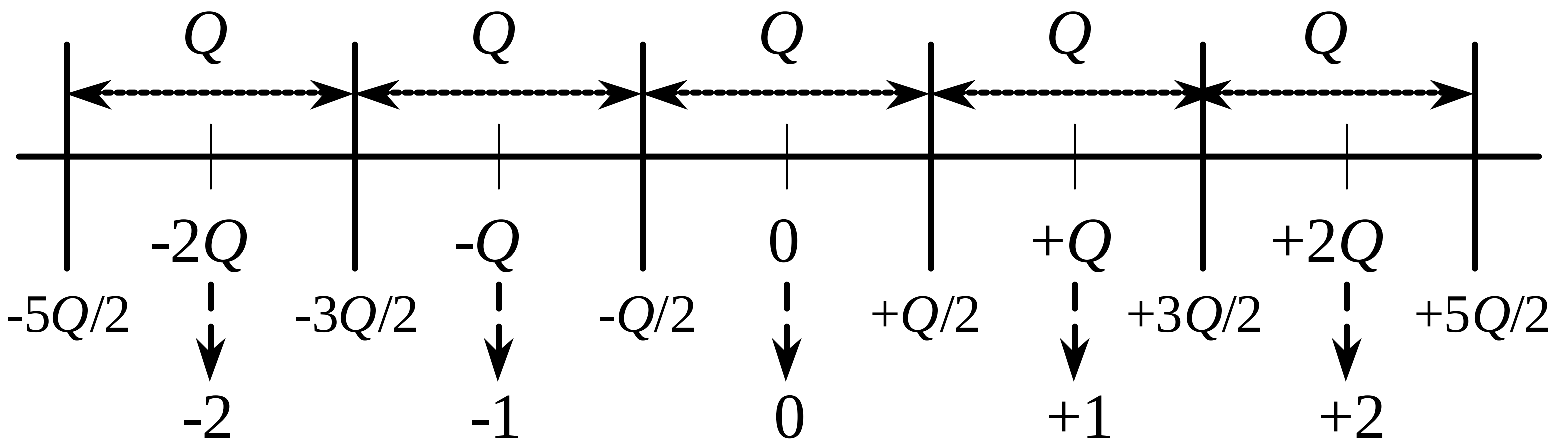


Quantization

Quantization

- A many-to-one mapping that reduces the number of possible signal values at the cost of introducing errors.
- The simplest form of quantization (also used in all the compression standards) is **scalar quantization** (SQ), where each signal value is individually quantized.
- The joint quantization of a block of signal values is called **vector quantization** (VQ). It has been theoretically shown that the performance of VQ can get arbitrarily close to the rate-distortion (R-D) bound by increasing the block size.
- In lossy compression schemes, quantization acts as a control knob for trading off image quality for bit rate (compression ratio).

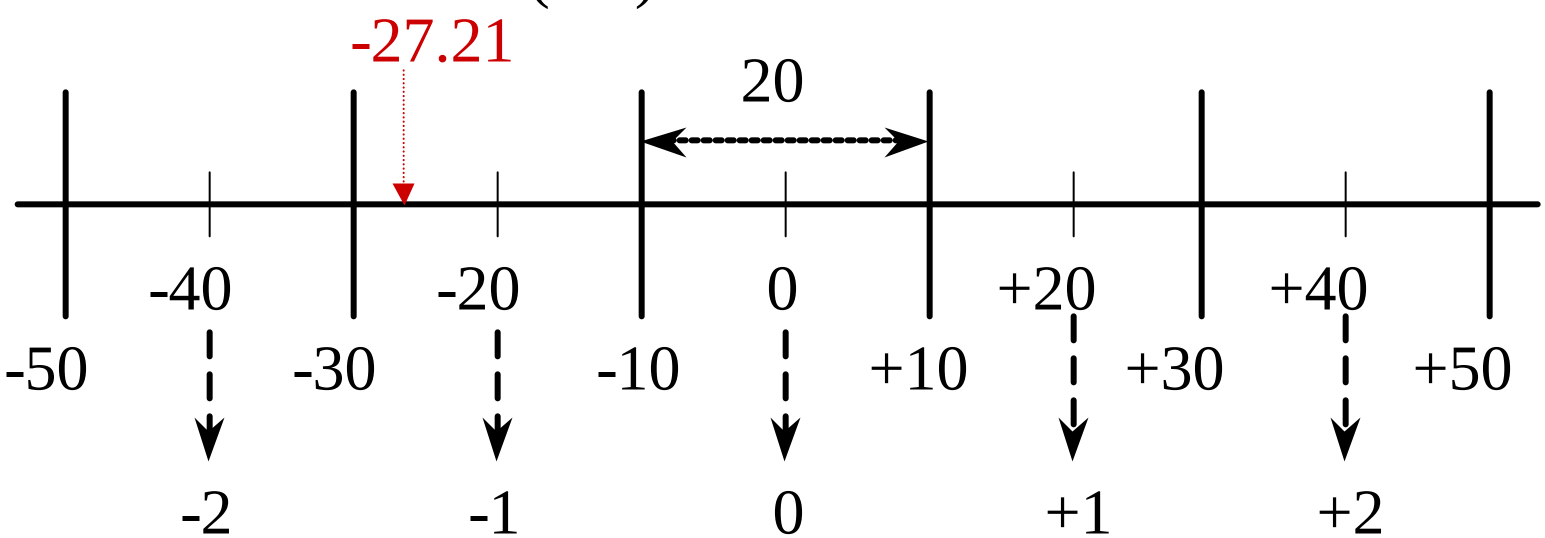
Uniform Threshold Quantizer (UTQ)



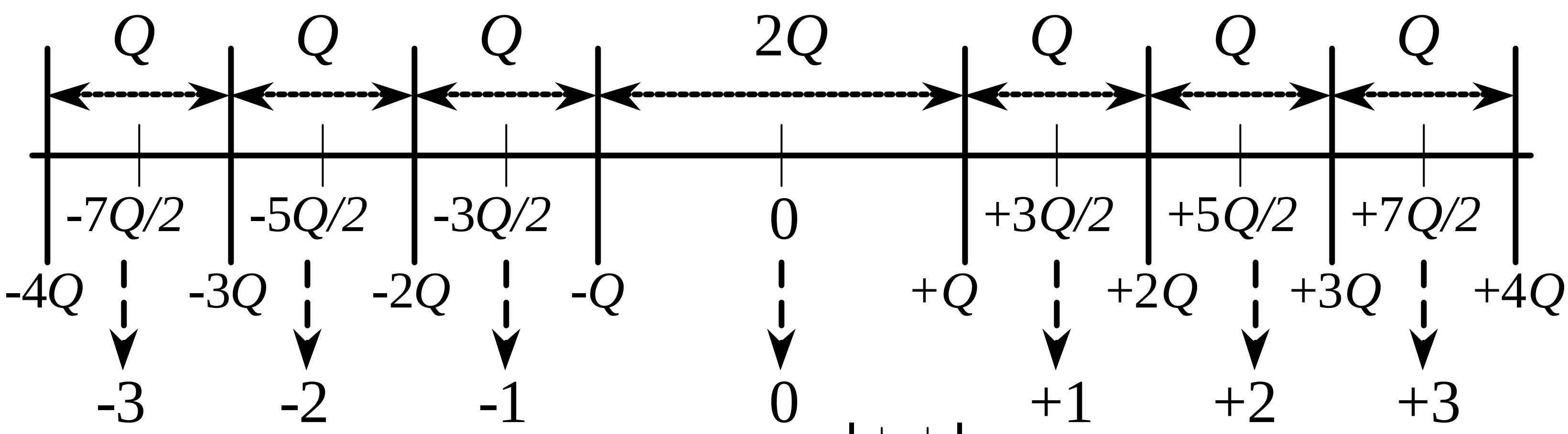
- In a **UTQ** quantizer, all bins have the same size. The bin size *Q* is called the quantizer **step size**. The quantization dequantization rule for a midpoint reconstruction is given by:
  - Quantization rule: $q = \mathrm{NINT}[y/Q]$
  - Dequantization rule: $z = q * Q$
  - Where *y* is the input signal, *q* is the resulting quantizer index, *z* is the **reconstructed** (quantized) value, and the NINT operation denotes rounding to the nearest integer.

Example: UTQ

- Quantization: encoder input value = -27.21
  - Scale by the step size → (-27.21)/(20) = -1.3605
  - Round to the nearest integer to get quantizer index = -1
- Dequantization: decoder received index = -1, step size = 20
  - Multiply quantizer index by step size → -1 x 20 = -20
  - Error = -27.21 – (-20) = -7.21



Uniform Threshold Quantizer with Deadzone



- Quantization rule: $z = \text{sign}\left\lfloor \frac{|y|}{Q} \right\rfloor$
- Dequantization rule: $z = (q + r * \text{sign}(q)) * Q$

where y is the input signal, q is the quantizer index, z is the reconstructed signal value, sign(x) is sign of x, $\lfloor x \rfloor$ denotes the largest integer smaller than x, and r is the reconstruction bias (r = 0.5 corresponds to midpoint reconstruction).

Symbol Modeling And Encoding

Symbol Modeling and Encoding

- Symbol modeling and encoding involves the process of defining a statistical model for the symbols to be encoded (e.g., quantizer output levels or indices) and assigning a binary codeword to each possible output symbol based on its statistics.
- The resulting code should be **uniquely decodable**, i.e., each string of input symbols should be mapped into a unique string of output binary symbols.
- Examples are fixed-length coding, **Huffman** coding, **Golomb-Rice** coding, **arithmetic** coding, **Lempel-Ziv-Welch** (LZW) coding.

Huffman Codes

Pixel Value	Probability	Code 1 Fixed	Code 2 Huffman
0	0.60	00	0
1	0.30	01	10
2	0.05	10	110
3	0.05	11	111

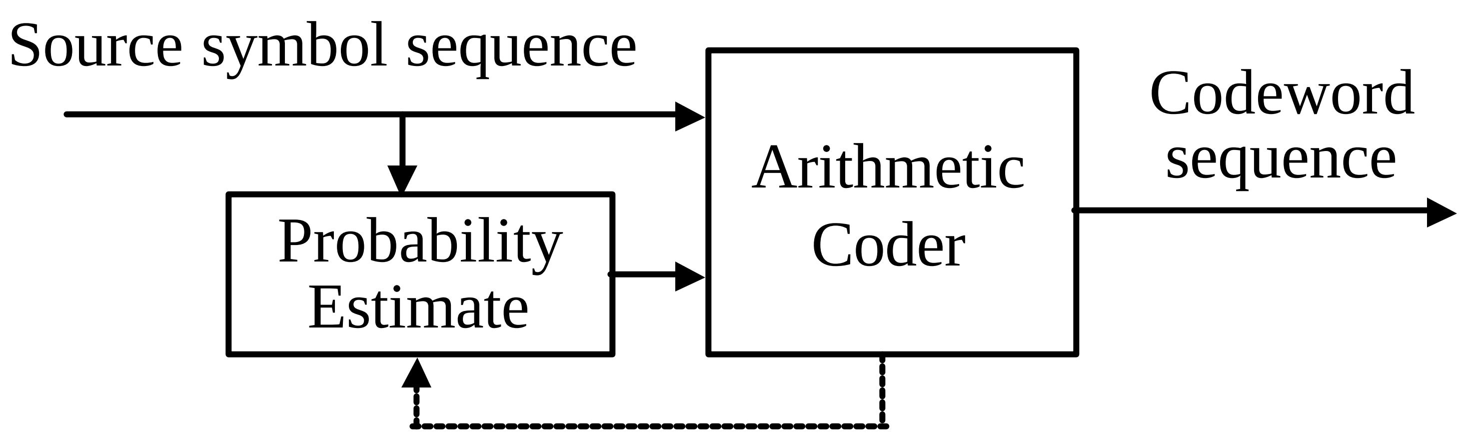
Example

Line 1	0	0	4	0	0	0	1	0	1	1
Code 1	00	00	11	00	00	00	01	00	01	01
Code 2	0	0	111	0	0	0	10	0	10	10
Line 2	0	0	3	0	0	0	1	0	1	1
Code 1	00	00	10	00	00	00	01	00	01	01
Code 2	0	0	110	0	0	0	10	0	10	10

- Average length of Code 1 = 2.0 bits/symbol.
- Average length of Code 2 = 1.5 bits/symbol.
- Code 2 is a prefix code, i.e., no codeword is a prefix of any other codeword (uniquely decodable)
- A Huffman code has an average length that is less than, or equal to, the average length of all other uniquely decodable codes for the same source and code alphabet.

Arithmetic Coding (AC)

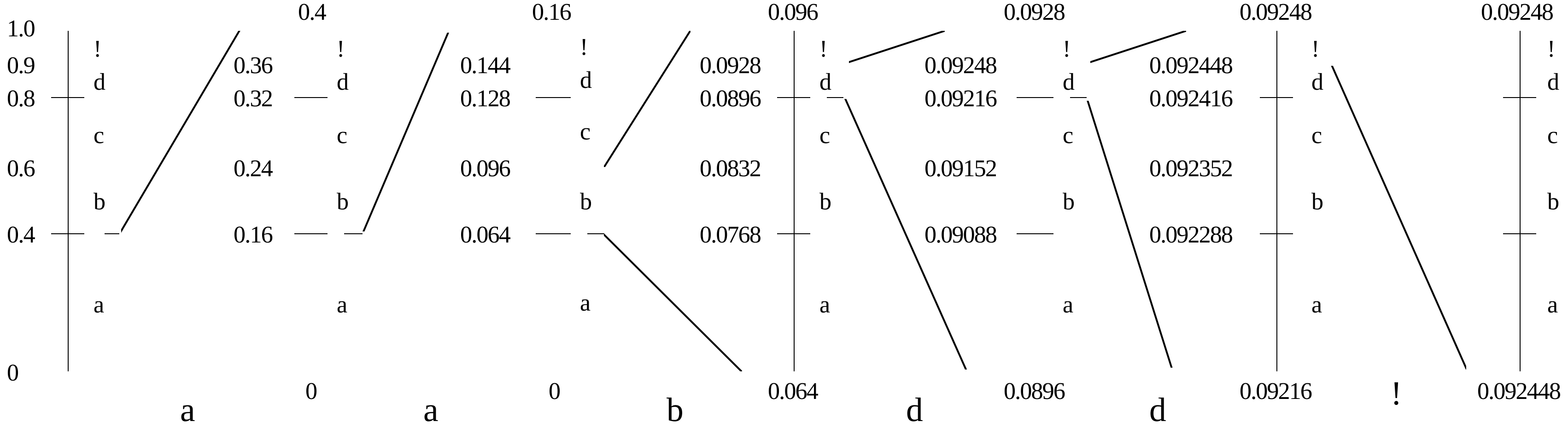
- An arithmetic coder accepts at its input the symbols in a source sequence along with their corresponding probability estimates, and produces at its output a code stream with a length equal to the combined ideal codelengths of the input symbols.
- Some implementations of arithmetic coding adaptively update the symbol probability estimate in each context as the symbols get encoded.
- Practical implementations of AC, such as the JBIG/JPEG **QM-Coder** or **MQ-Coder**



Arithmetic Coding

- Example m-ary (base 10, want base 2 and binary)
- Decoder stops with special symbol or by counting

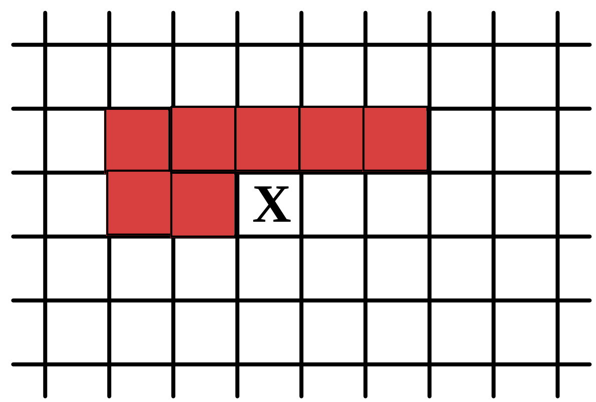
Symbol	Probability	CDF
a	0.4	[0.0, 0.4)
b	0.2	[0.4, 0.6)
c	0.2	[0.6, 0.8)
d	0.1	[0.8, 0.9)
!	0.1	[0.9, 1.0]



Conditioning Contexts

- In general, the probability of a sample having a certain value is influenced by the value of its neighbors. Thus, the symbol probabilities can be conditioned on the values of the symbols in a neighborhood surrounding them. For a given neighborhood configuration, each combination of the neighboring samples denotes a **conditioning context**.
- The **conditional entropy** of a correlated source can be significantly less than its zeroth-order entropy.

0	0	0	0	0	0	0	0
0	0	0	0	0	1	0	0
0	0	0	0	1	1	0	0
0	0	0	0	1	1	0	0
0	0	0	0	0	0	0	0



Example: Entropy of Lena MSB

Conditioning contexts can capture the redundancy in the image:

No conditioning contexts
Entropy = **1.0** bit/pixel

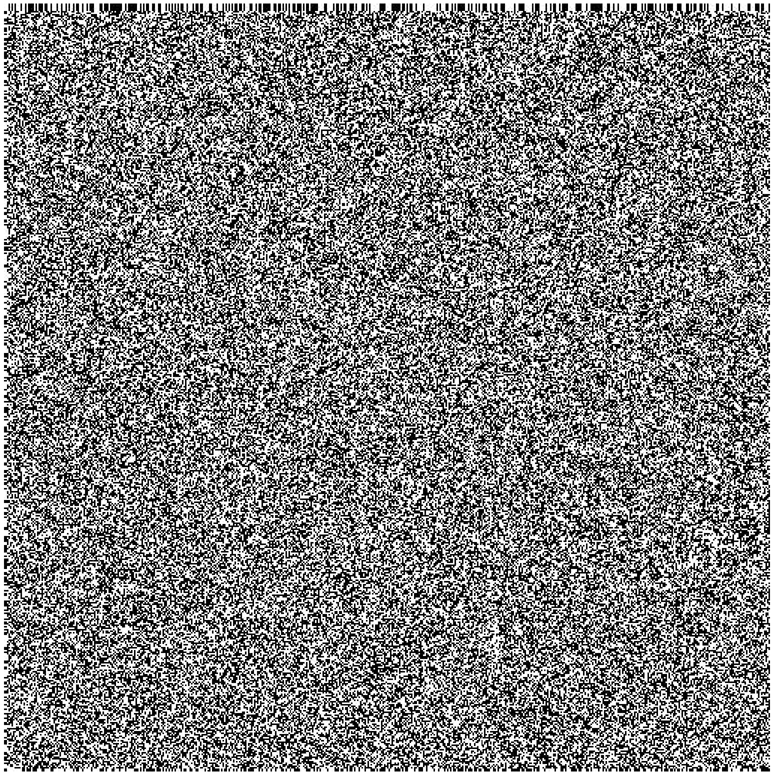
7-neighbor conditioning context
Entropy = **0.14** bits/pixel

Most significant bit plane

Example: Entropy of Lena LSB

No conditioning contexts
Entropy = **1.0** bit/pixel

7-neighbor conditioning context
Entropy = **1.0** bits/pixel



Least significant bit plane

Rate Controller

- A rate controller is used when an exact compression rate or image throughput is desired (e.g., DDS 1.3 DCT).
- The rate controller changes the amount of quantization dependent on the output bit rate and the desired bit rate.
  - The quantization is greater (i.e., bin size gets larger) when too many bits are coming out of the symbol encoder.
  - The quantization is reduced when too few bits are coming out of the symbol encoder.
- The rate control can be performed single-pass (the quantization step size changes as a function of location in the image) or multiple-pass (quantization step size is usually consistent throughout the image, tile or block).

Color and Multiple Component Transform

Color Image Representation

- Color image components are highly correlated due to:
  - Overlapping spectral responses of the sensors
  - Smooth spectral distribution of surfaces and illuminants
- The RGB color values are often transformed into a new set of values called **luminance** and **chrominance** (such as YCrCb, or YIQ), such that:
  - The transformed components are less correlated (reduced redundancy), and,
  - The sensitivity variations of the human visual system (irrelevancy) can be taken into account, e.g., the chromatic components may be subsampled or compressed more aggressively.

YC_bC_r Color Space

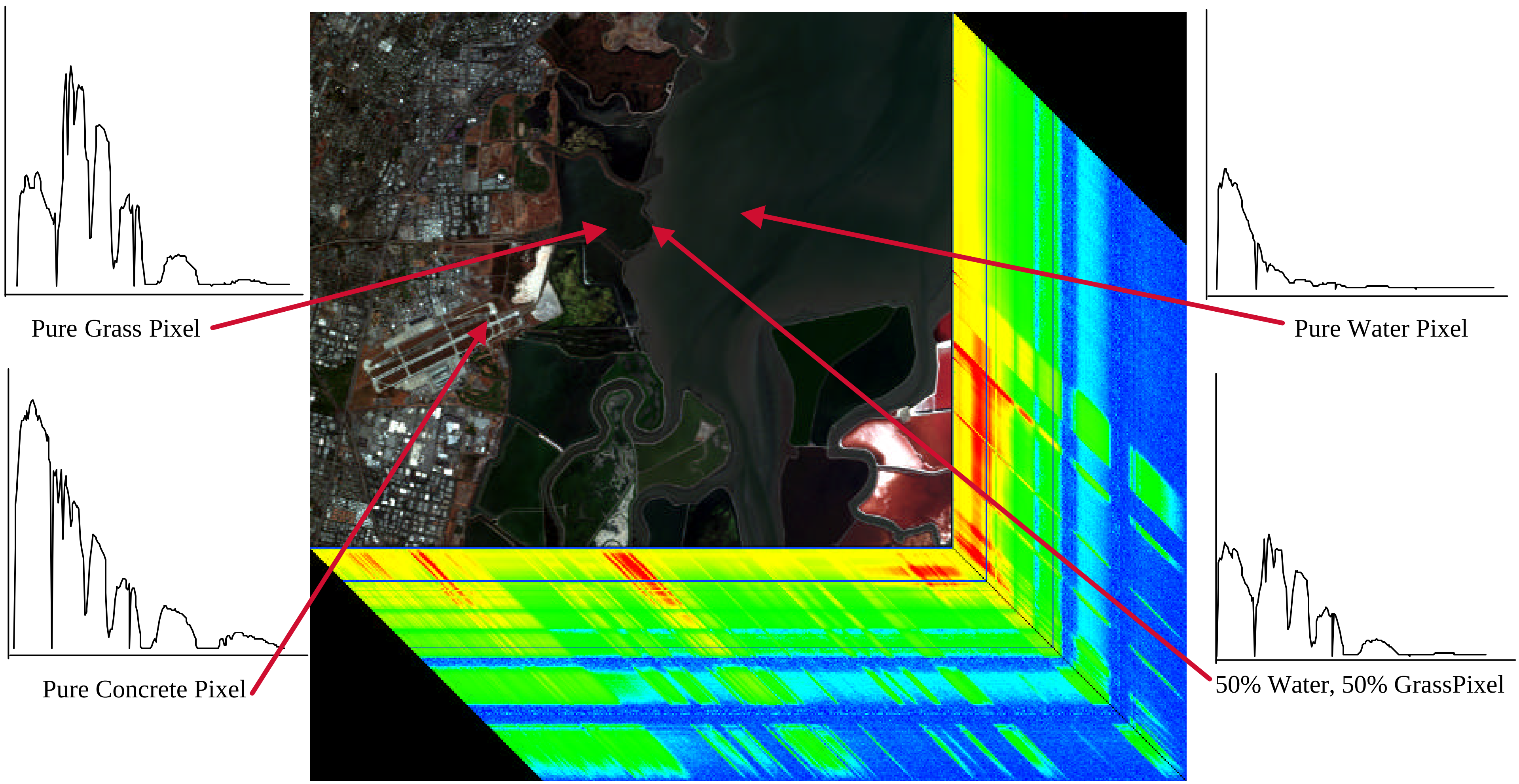
This is the most commonly used color coordinate system for the representation of image and video signals:

$$Y = 0.299(R - G) + G + 0.114(B - G)$$
$$C_b = 0.564(B - Y) \quad \text{and} \quad C_r = 0.713(R - Y)$$

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

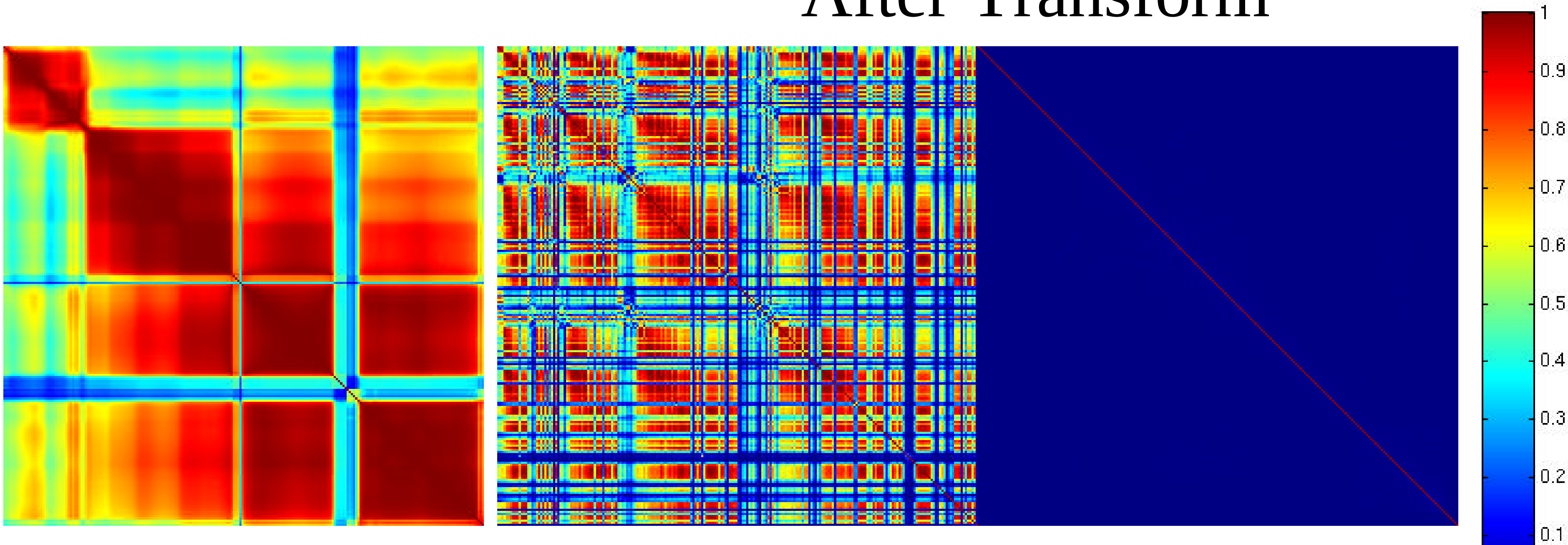
$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.0 & 0.0 & 1.4021 \\ 1.0 & -0.3441 & -0.7142 \\ 1.0 & 1.7718 & 0.0 \end{bmatrix} \begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix}$$

Hyperspectral Information Cube (AVIRIS)



Spectral Correlation (AVIRIS)

After Transform



Original | Spectral Wavelet | KLT

How To Choose a Compression Algorithm

Standards

Requirements

Choosing a Compression Algorithm

- Q: What is the best compression technique?
- A: It depends on the application!
- Some factors to consider:
  - Image quality (lossless, visually lossless, visually lossy, acceptable loss)
  - Operational bit rate (transmission rate vs. image size/number requirements)
    - Constant bit rate(per block) vs. fixed bit rate(per pixel) vs. constant quantization
  - Computational complexity
  - Channel error tolerance
  - Encoder/decoder asymmetry
  - Artifacts (blocking, noise, edge blur)
  - System compatibility and compression standards
  - Input image characteristics
    - Data type and previous processing (sharpening, compression)
  - Output image applications
  - Spatial Accuracy

Digital Image Compression Standards

- Facilitate the exchange of compressed image data between various devices, applications and users.
- Permit common hardware/software to be used for a wide range of products, thus lowering costs and shortening development time.
- Several levels of standards:
  - Specification used in limited-access world
    - 1.3 DCT, 2.3 DCT, 4.3 DPCM
  - Military Standard used in DoD community
    - MIL-STD-188-198A NITFS JPEG DCT, NITFS Vector Quantization
  - International standards used in the commercial world
    - ISO/IEC 10918-1 (JPEG)
      - Very broad tool box; not all JPEG algorithms are the same

Commercial Image Compression Standards

- Binary (bi-level) images:
  - Group 3 & 4 facsimile ('80); JBIG ('94); JBIG2 ('01)
- Continuous-tone still images:
  - JPEG ('92); JPEG2000 (Part 1 IS '00, Part 2 IS '01, other parts ongoing) – www.jpeg.org
- Image sequences (moving pictures):
  - H.261 ('90); H.263 ('95); H.263+ ('97); H.26L (merged with MPEG4 AVC)
  - MPEG1 ('94); MPEG2 ('95); MPEG4 ('99, AVC still ongoing) - http://mpeg.telecomitalialab.com/

Standards Background

- 4.3 DPCM
  - Developed for visually lossless, rate-controlled simple compression for storage and transmission
  - Old technology, current technology can significantly outperform
- 1.3 DCT/2.3 DCT
  - Significant development effort to produce a high quality (0.2/0.1 NIIRS loss) at low bit rates (1.3 BPP/2.3 BPP).
  - Old technology, still very competitive but not very flexible
- JPEG DCT/NITFS JPEG DCT
  - Developed as a commercial standard to run on commercial PCs (386s) and commercially viable hardware.
  - NITFS/DoD adopted because of quality, flexibility and COTS products
- Vector Quantization
  - Developed to compress maps with very fast decompression.
  - Used by NIMA to put maps and imagery out on CD.
- NIMA Method 4
  - Developed to achieve dissemination to warfighters with very low bandwidth communication lines

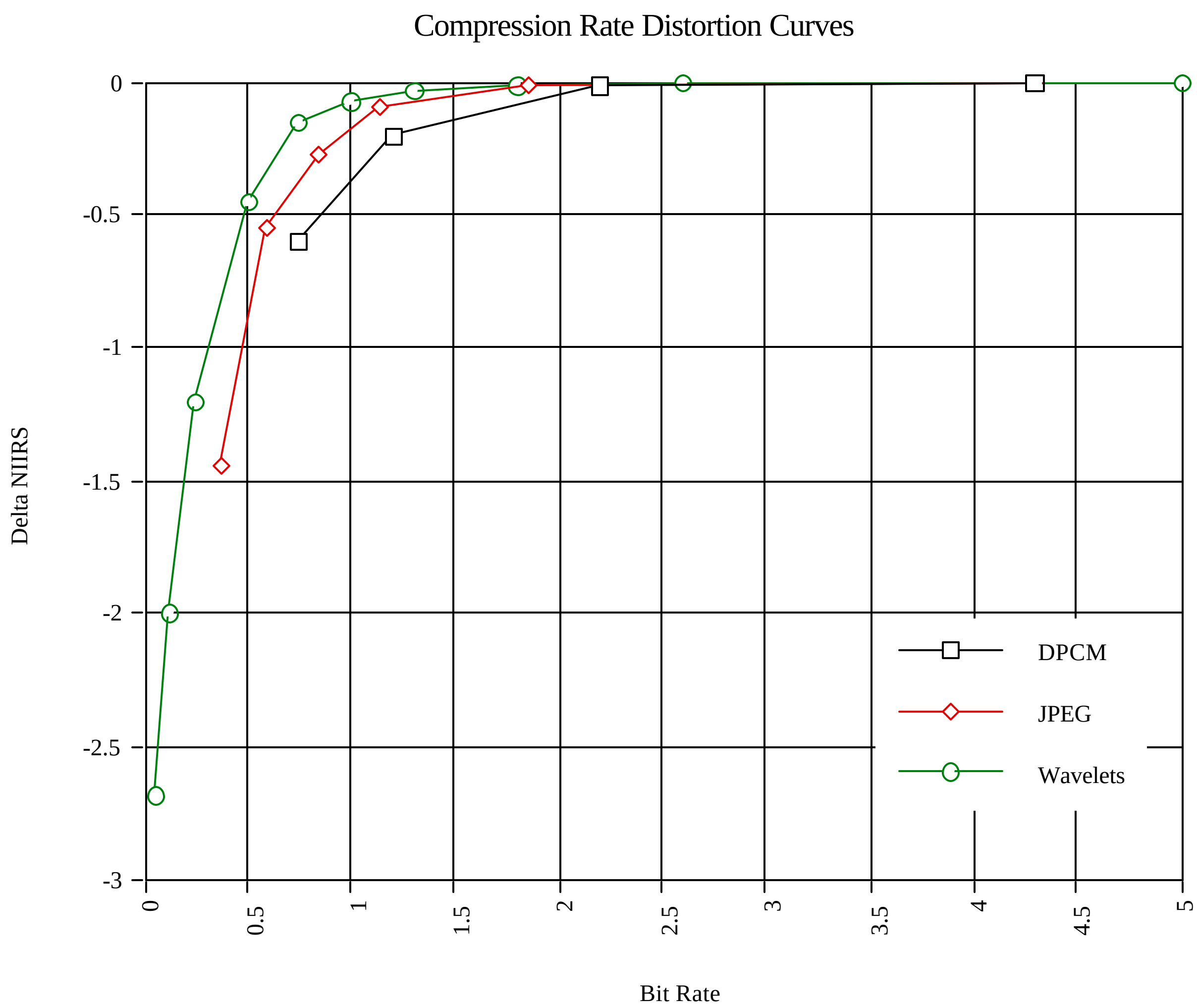
Current Requirements

- 4.3 DPCM
  - 0.0 NIIRS loss, 2:1 or better compression, fast decompression
- 2.3 DCT
  - 0.1 NIIRS loss, 3:1 or better compression, spatial accuracy
- 1.3 DCT
  - 0.2 NIIRS loss or less, 1.3 bpp or less, robust to channel errors
- NITFS/NIMA VQ
  - Fast decompression, variable compression, robust to channel errors
- NITFS JPEG DCT
  - 0.5 NIIRS loss at 8:1 compression, 2.0 min. decompression time Variable compression ratios, robust to channel errors
- NIMA Method 4
  - High compression ratios with minimal image quality loss

Compression Optimization

- Each compression algorithm has several parameters that can be modified to improve the quality, increase the compression ratio (at same quality) or reduce artifacts.
  - For example, JPEG optimization can give a 5% to 15% gain in compression with proper optimization of the quantization and Huffman tables or a 0.5 NIIRS improvement at the same compression rate
  - Parameters are optimized for the characteristics of the image and/or the requirements of the compression applications
    - Optimization is common for a class of imagery or image characteristics
      - Color, panchromatic, IR, SAR, noisy, graphic
    - Optimization is also common for a desired bit rate (1.3 bpp, 2.3 bpp)
      - Quantization tables, Huffman tables
  - Parameters can be modified to reduce identified artifacts which may be the interaction between the compression algorithm, the image characteristics, post processing and the display process.

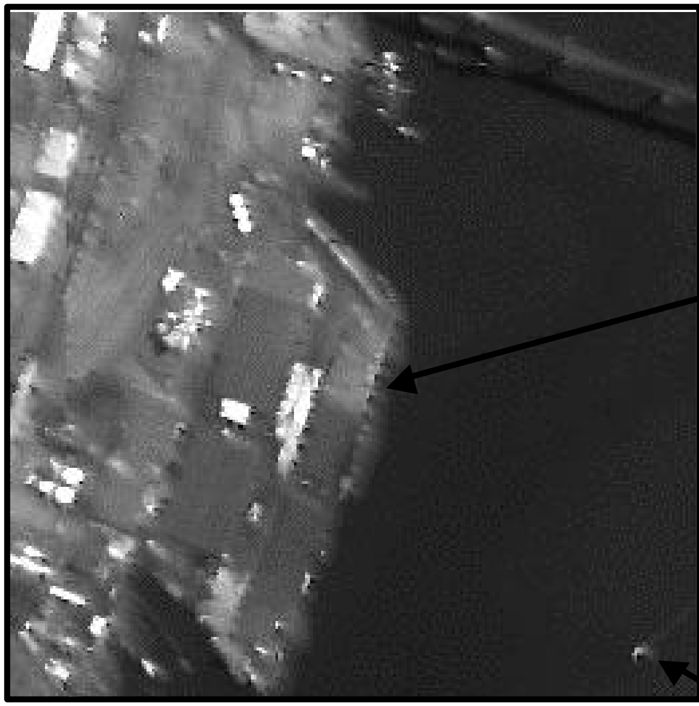
Compression Rate Distortion



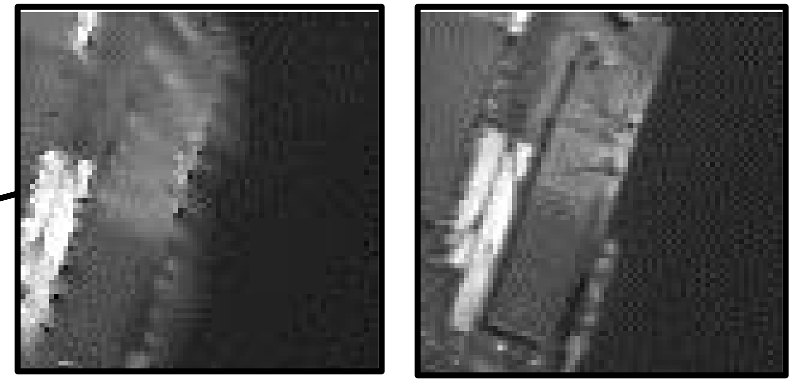
Compression Artifacts

- Artifacts of compression are viewable when:
  - The compression ratio is pushed beyond the normal working environment of the given compression algorithm, or
  - the image is processed beyond the "normal" range of enhancements (i.e., sharpen, sharpen-more, DRA, TTC)
- Common artifacts include;
  - DPCM
    - Slope overload, water-fall artifact
  - DCT
    - Blocking, ringing around edges, DCT basis functions
  - Wavelets
    - False texture, reduction in resolution, ringing
  - VQ
    - Blocking, contouring

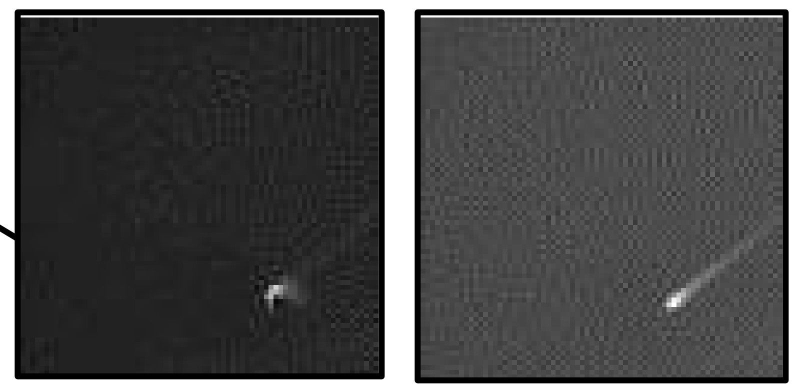
DPCM Example (1.8 bpp)



- Artifacts include;
  - Slope overload
  - Water-fall artifact

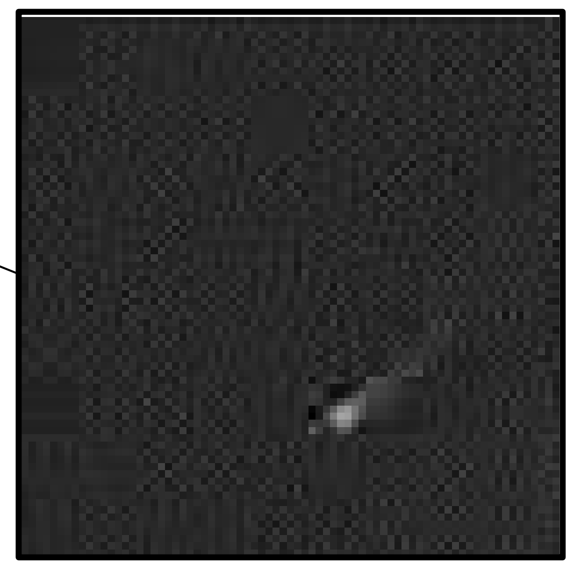


DPCM Original



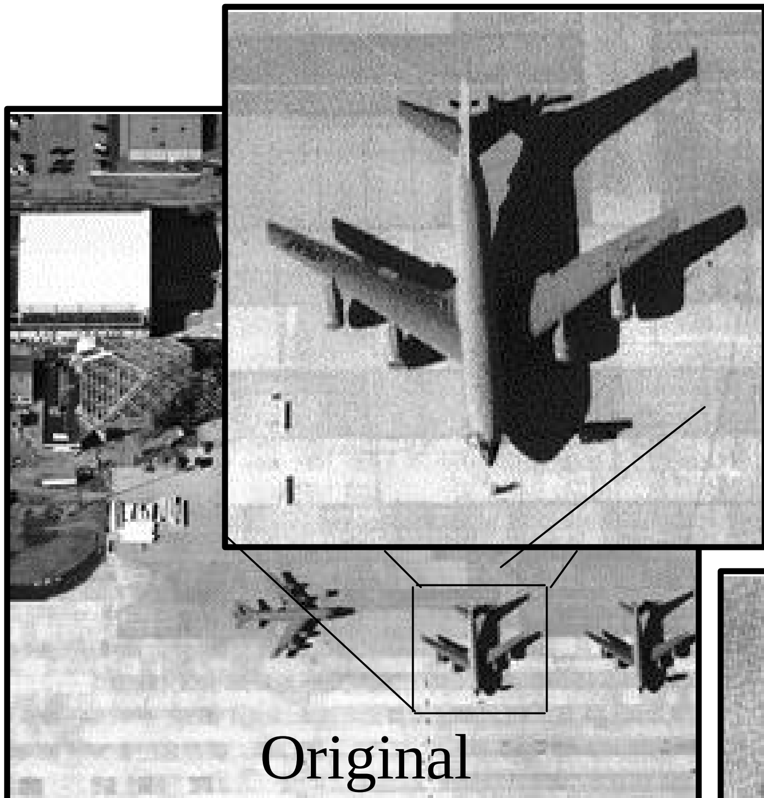
DPCM Original

DCT Example (JPEG @ 0.4 bpp)

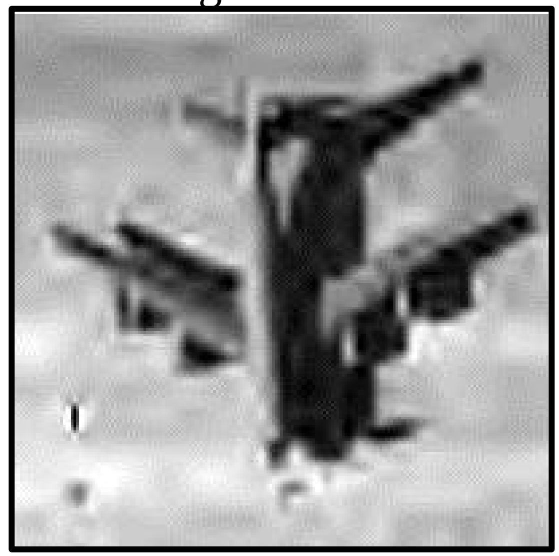


- Artifacts include;
  - Blocking
  - Ringing around edges
  - DCT basis functions

Wavelet Example (0.0625)



Algorithm 1



Algorithm 2

Algorithm 3

- Artifacts include;
  - False texture
  - Reduction in resolution
  - Ringing

Channel Errors

- Problems from channel errors are hard to characterize for each algorithm
- Several factors affect the image quality when a channel error is occurred
  - Variable length encoder vs. fixed length encoders
    - A channel error in a variable length encoder will propagate until the encoder resyncs or there is a restart interval
    - A channel error in a fixed length encoder only affects that value
  - Prediction/transform technique
    - Any incorrect value is propagated to surrounding value depending on the prediction or transform technique
      - Only the block of a given DCT is affected by an error in the AC components
      - Error is propagated from the error pixel to the lower and right for a DPCM
      - Depending on the level of the wavelet the error is propagated to the surrounding 2N by 2N pixels (N is the level to error occurred)

Overcoming Channel Errors

- Protection to channel errors
  - Restart markers
    - Restart markers are used to restart the algorithm to stop the propagation of any error that may have occurred before
  - Error Dection And Correction (EDAC)
    - Forward Error Correction (FEC)
      - Will correct errors automatically
    - Error Detection
      - Can detect errors for retransmission of data
  - Re-send data (the simplest of all methods)
    - Re-send data that is bad 2-3 times and make decision (2/3 rule)
  - These techniques can be used on the entire data or data that is determined to be critical
    - For example, the DC component, Huffman tables, quantization tables of JPEG DCT

Image Sequence Compression (Video)

- Image sequences (neighboring frames) are often highly correlated, particularly if object motion is taken into account (motion compensation)
- Motion-compensated frame differencing can be used very effectively to reduce redundant information in sequences.
- Finding corresponding points between frames (i.e., motion estimation) can be difficult because of occlusion, noise, illumination changes, etc.
- Motion vectors (x,y-displacements) are sent to the receiver to indicate corresponding points; these vectors are usually computed over blocks of pixels (e.g., 16x16) to minimize overhead.

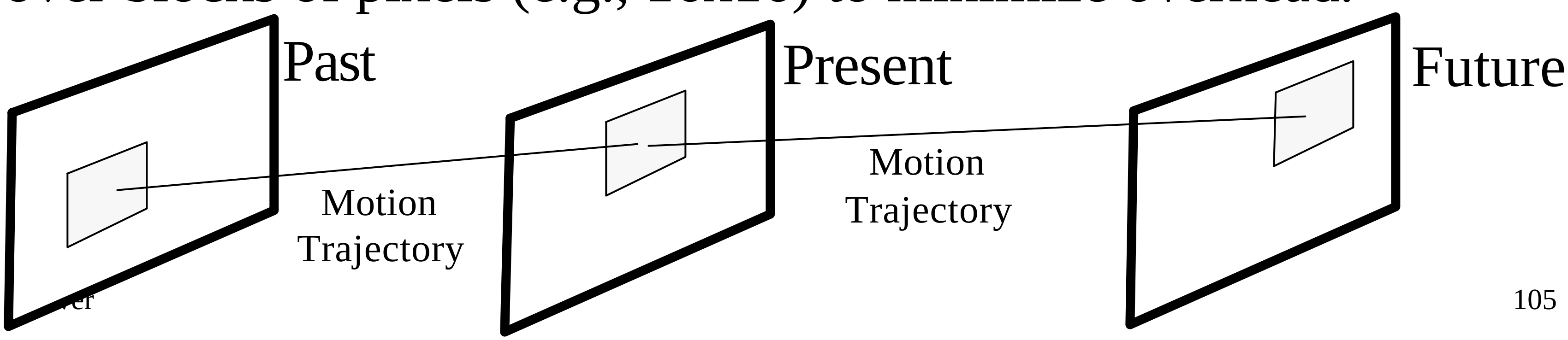
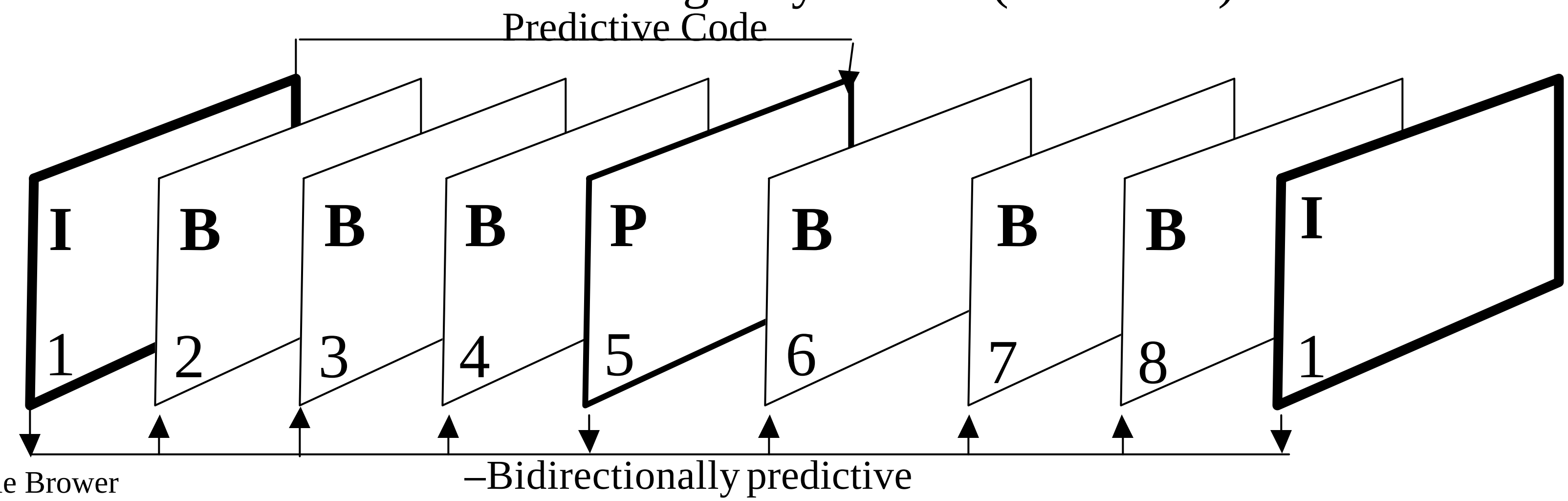


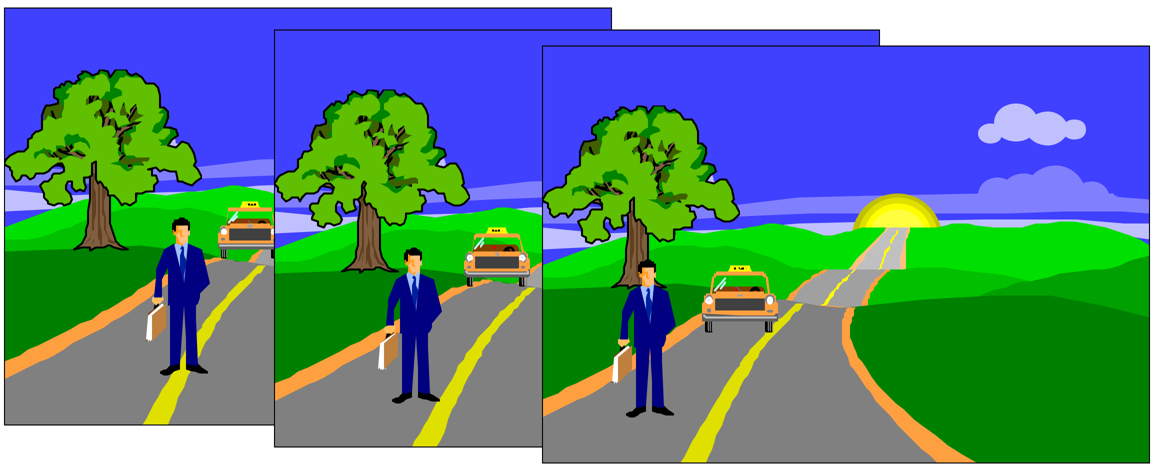
Image Sequence Compression (Video)

- The MPEG system specifies three types of frames within a sequence:
  - **Intra-coded picture** (I-frame): Coded independently from all other frames in the sequence. Uses the most number of bits.
  - **Predictive-coded picture** (P-frame): Coded based on a prediction from a past I- or P-frame. Uses less bits than an I-frame.
  - **Bidirectionally predictive coded picture** (B-frame): Coded based on a prediction from a past and/or future I- or P-frame(s). Uses the least number of bits and cannot be used as a reference for prediction.
  - Each frame is encoded using 8-by-8 DCT (JPEG like)

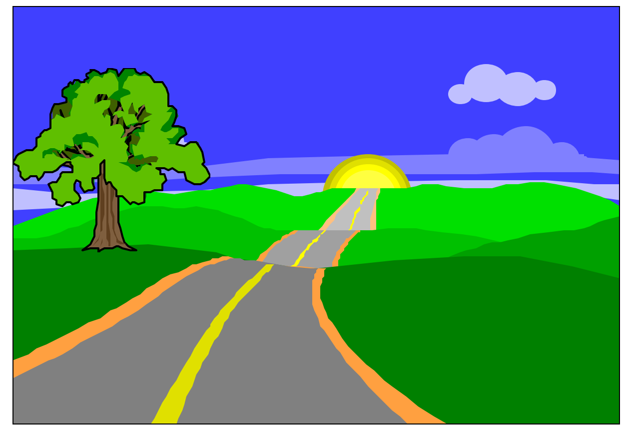


MPEG-7

- Object based motion compression
  - Separate objects (background, object 1, object 2)
  - Compress each object separate
    - Send updates to objects not background



Original Scene



Background

Object 1

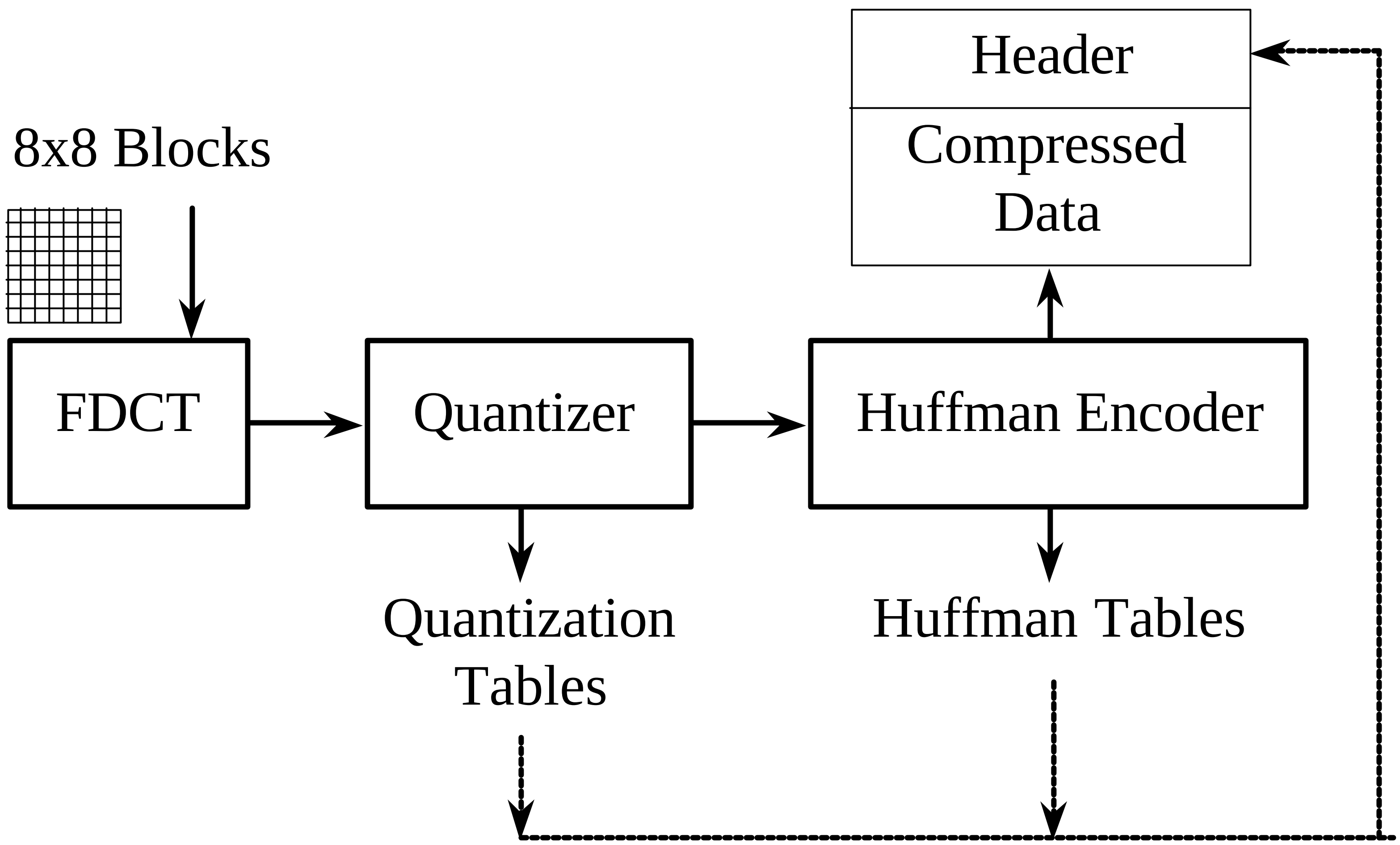
Object 2

The Existing DCT-Based JPEG Standard

What Is JPEG?

- The JPEG (**Joint Photographic Experts Group**) committee, formed in 1986, has been chartered with the
  - "*Digital compression and coding of continuous-tone still images*"
- Joint between ISO and ITU-T
- Has developed standards for the lossy, lossless, and nearly lossless of still images in the past decade
- Website: **www.jpeg.org**
- The JPEG committee has published the following standards:
  - ISO/IEC 10918-1 | ITU-T Rec. T.81 : *Requirements and guidelines*
  - ISO/IEC 10918-2 | ITU-T Rec. T.83 : *Compliance testing*
  - ISO/IEC 10918-3 | ITU-T Rec. T.84: *Extensions*
  - ISO/IEC 10918-4 | ITU-T Rec. T.86: *Registration of JPEG Parameters, Profiles, Tags, Color Spaces, APPn Markers Compression Types, and Registration Authorities (REGAUT)*

JPEG Encoder Block Diagram



JPEG Decoder Block Diagram

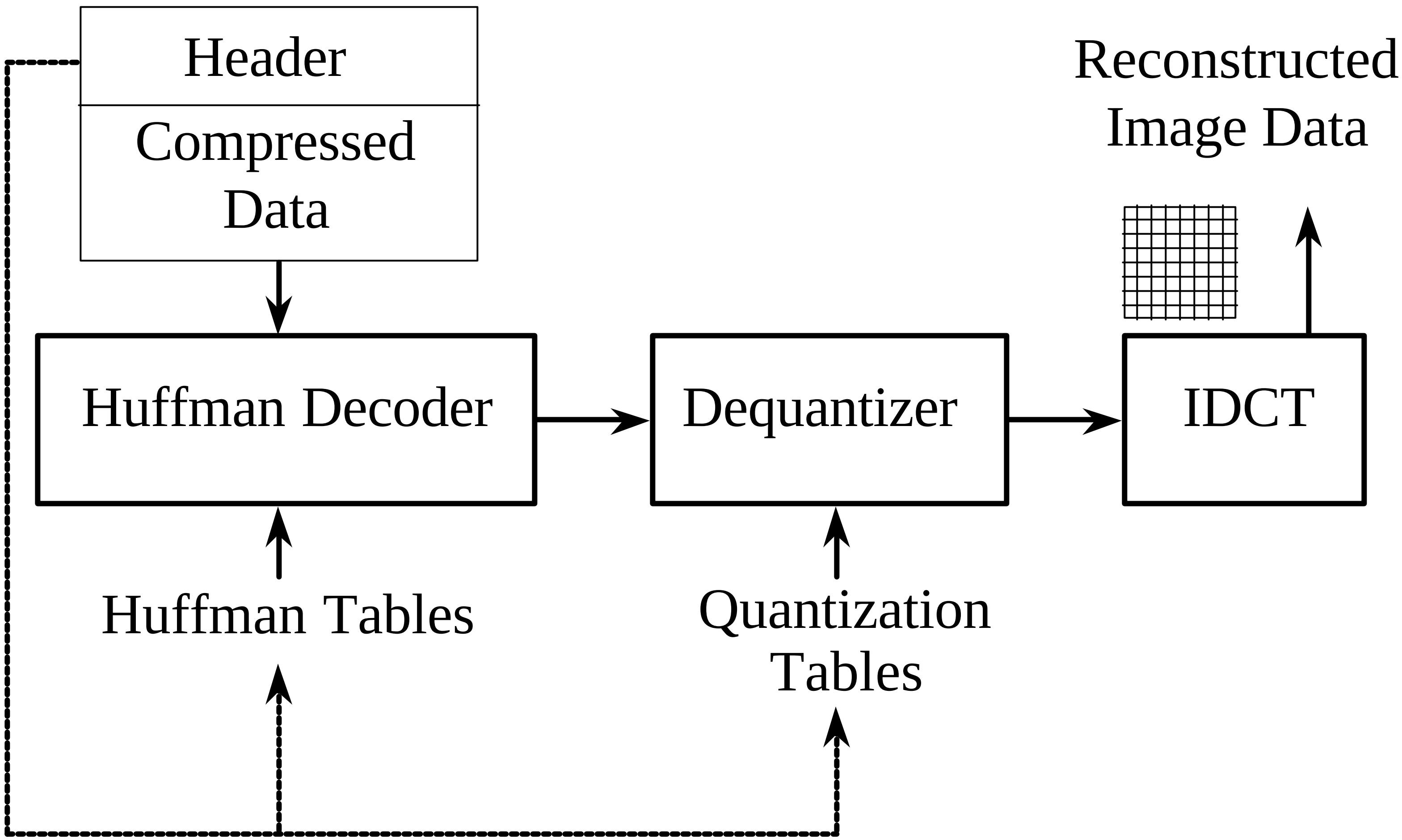
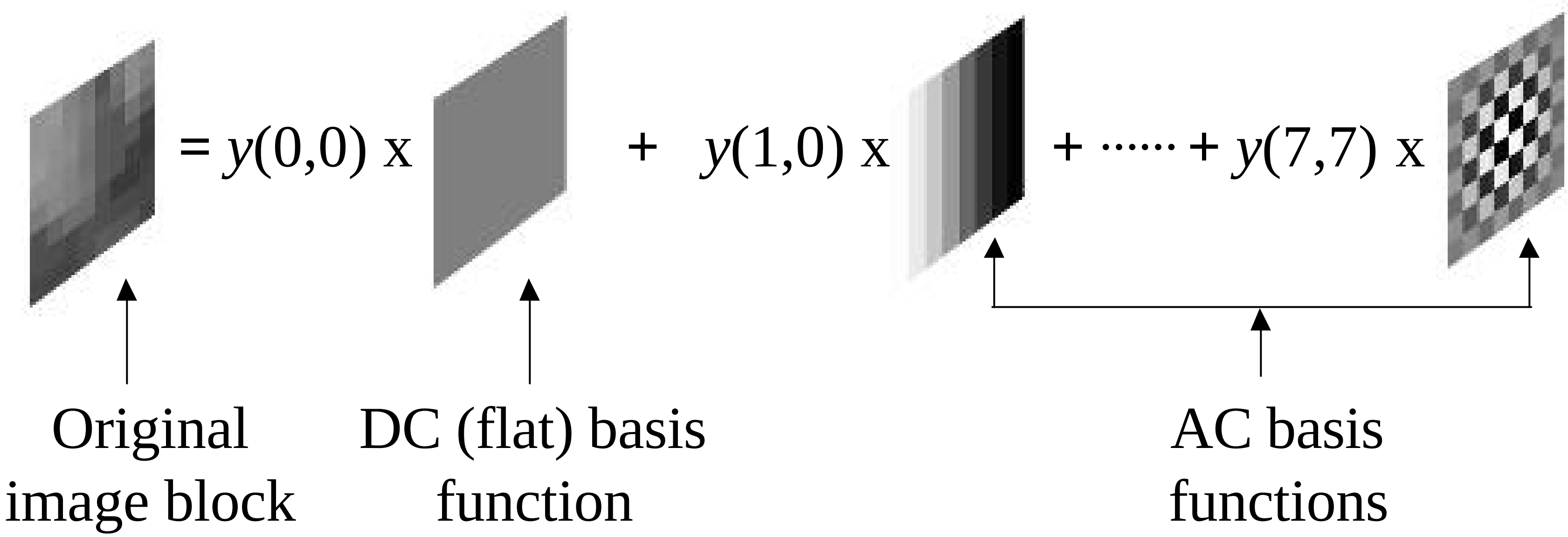


Image Representation with DCT

- DCT coefficients can be viewed as weighting functions that, when applied to the 64 cosine basis functions of various spatial frequencies (8 x 8 templates), will reconstruct the original block.



JPEG DCT Example

Original Image

139	144	149	153	155	155	155	155
144	151	153	156	159	156	156	156
150	155	160	163	158	156	156	156
159	161	162	160	160	159	159	159
159	160	161	162	162	155	155	155
161	161	161	161	160	157	157	157
162	162	161	163	162	157	157	157
162	162	161	161	163	158	158	158

8 BPP
64 pixels
512 bits

DCT Transformed Image

235.6	-1.0	-12.1	-5.2	2.1	-1.7	-2.7	1.3
-22.6	-17.5	-6.2	-3.2	-2.9	-0.1	0.4	-1.2
-10.9	-9.3	-1.6	1.5	0.2	-0.9	-0.6	-0.1
-7.1	-1.9	0.2	1.5	0.9	-0.1	0.0	0.3
-0.6	-0.8	1.5	1.6	-0.1	-0.7	0.6	1.3
1.8	-0.2	1.6	-0.3	-0.8	1.5	1.0	-1.0
-1.3	-0.4	-0.3	-1.5	-0.5	1.7	1.1	-0.8
-2.6	1.6	-3.8	-1.8	1.9	1.2	-0.6	-0.4

Quantized/ Scaled Transformed Data

15	0	-1	0	0	0	0	0
-2	-1	0	0	0	0	0	0
-1	-1	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

DCT Coefficient Quantization

- Each DCT coefficient is uniformly quantized with a quantization step that is taken from a user-defined **quantization table** (q-table or normalization matrix), characterized by 64, 1-byte elements.
- The quality and compression ratio of an encoded image can be varied by changing the q-table elements (usually by scaling up or down the values of an initial q-table).
- The q-table is often designed according to the perceptual importance of the DCT coefficients (e.g., by using the HVS CSF data) under the intended viewing conditions.
- For the baseline system, in order to meet the needs of the various color components, up to four different quantization tables are allowed.

Example of Luminance Quantization Table

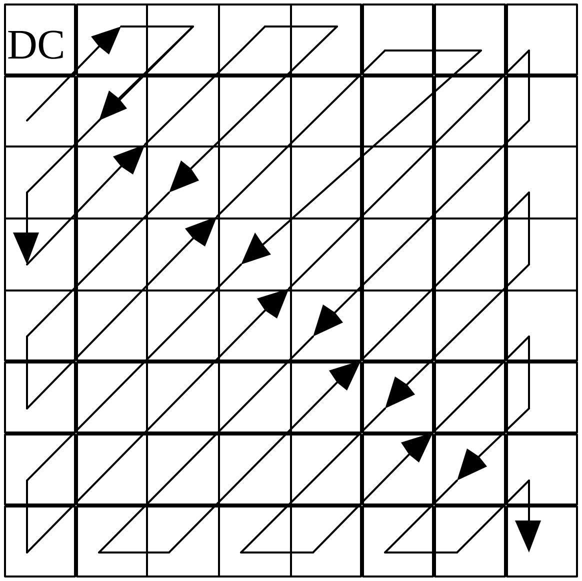
The JPEG committee has listed the following luminance quantization table as an example in Annex K (informative) of the IS. It is obtained by measuring the DCT coefficient "visibility threshold" using CCIR-601 images and display, at a distance of six picture-heights away.

$Q_L(u,v) =$

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

JPEG DCT Example

- After Quantization, the DCT is separated into a DC coefficient and AC coefficients, which are reordered into a 1-D format using a zigzag pattern in order to create long runs of zero-valued coefficients.
  - The DC coefficient is directly correlated to the mean of the 8-by-8 block (upper-left corner).
    - All DC coefficients are combined into a separate bit stream.
  - The AC coefficients are the values of the cosine basis functions (all other values).



JPEG DCT Example

DC	AC								
15	0	-2	-1	-1	-1	0	0	-1	EOB

- The DC coefficient is encoded using Huffman encoded 1D-DPCM
- The AC coefficients are encoded using Huffman coding on magnitude/runlength pairs (magnitude of a nonzero AC coefficient plus runlength of zero-valued AC coefficients that precede it).
- The end-of-block (EOB) symbol indicates that all remaining coefficients in the zigzag scan are zero. This allows many coefficients to be encoded with only a single symbol.

JPEG DCT Example

Dequantized DCT Coefficients

240	0	-10	0	0	0	0	0
-24	-12	0	0	0	0	0	0
-14	-13	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

Reconstructed Image

144	146	149	152	154	156	156	156
148	150	152	154	156	156	156	156
155	156	157	158	158	157	156	155
160	161	161	162	161	159	157	155
163	163	164	163	162	160	158	156
163	163	164	164	162	160	158	157
160	161	162	162	162	161	159	158
158	159	161	161	162	161	159	158

Error Image

-5	-2	0	1	1	-1	-1	-1
-4	1	1	2	3	0	0	0
-5	-1	3	5	0	-1	0	1
-1	0	1	-2	-1	0	2	4
-4	-3	-3	-1	0	-5	-3	-1
-2	-2	-3	-3	-2	-3	-1	0
2	1	-1	1	0	-4	-2	-1
4	3	0	0	1	-3	-1	0

RMSE= 2.26

The Emerging JPEG2000 Standard

JPEG-DCT Pros and Cons

- <u>Advantages</u>
  - Memory efficient
  - Low complexity
  - Compression efficiency
  - Visual model utilization
  - Robustness
- <u>Disadvantages</u>
  - Single resolution
  - Single quality
  - No target bit rate
  - No lossless capability
  - No tiling
  - No region of interest
  - Blocking artifacts
  - Poor error resilience

JPEG2000 Objectives

- Advanced standardized image coding system to serve applications into the next millenium
- Address areas where current standards fail to produce the best quality or performance
- Provide capabilities to markets that currently do not use compression
- Provide an open system approach to imaging applications

JPEG2000 Compression Standard Status

- The standard only specifies a decoder and a bitstream syntax and is issued in several parts:
  - **Part 1**: Specifies the minimum compliant decoder (e.g., a decoder that is expected to satisfy 80% of applications). International Standard (IS) approved 12/30/00 and published by the ISO. Intended as royalty and license-fee free, however, NOT patent-free.
  - **Part 2**: Describes optional features and value added extensions. International Standard (IS) was approved 10/1/01. Some technology is covered by IPR.
  - **Part 3**: Motion JPEG 2000 with file format from MPEG 4. IS approved in 2001.

JPEG2000 Compression Standard Status

- **Part 4**: Compliance testing procedures. Compliance test image and procedures are very important for the promotion of "compliant" standards and interoperability.
- **Part 5**: Reference software. Two versions of reference software International Standard have been approved:
  - JAVA - http://jj2000.epfl.ch
  - C - http://www.ece.ubc.ca/~mdadams/jasper/
- **Part 6**: Compound document. Being developed to support compound documents (text, graphics, and images) using the Mixed Raster Content (MRC) defined in ISO 16458. Currently at FCD.
- **Part 7:** abandoned.

New Work Items

- **Part 8:** Security **JPSEC**. Provides tools and solutions in terms of specifications that allow applications to generate, consume, and exchange secure JPEG 2000 bitstreams.
- **Part 9:** Interactivity tools, APIs and Protocols. Supports user interaction with JPEG 2000 images by providing APIs whereby applications could exploit JPEG features, and supports protocols for remote interaction over networks.
- **Part 10:** 3-D and floating point data**.** Relates to compression and decompression of volumetric data.
- **Part 11:** **JPWL** (wireless applications)
- **Part 12:** ISO Base Media File Format

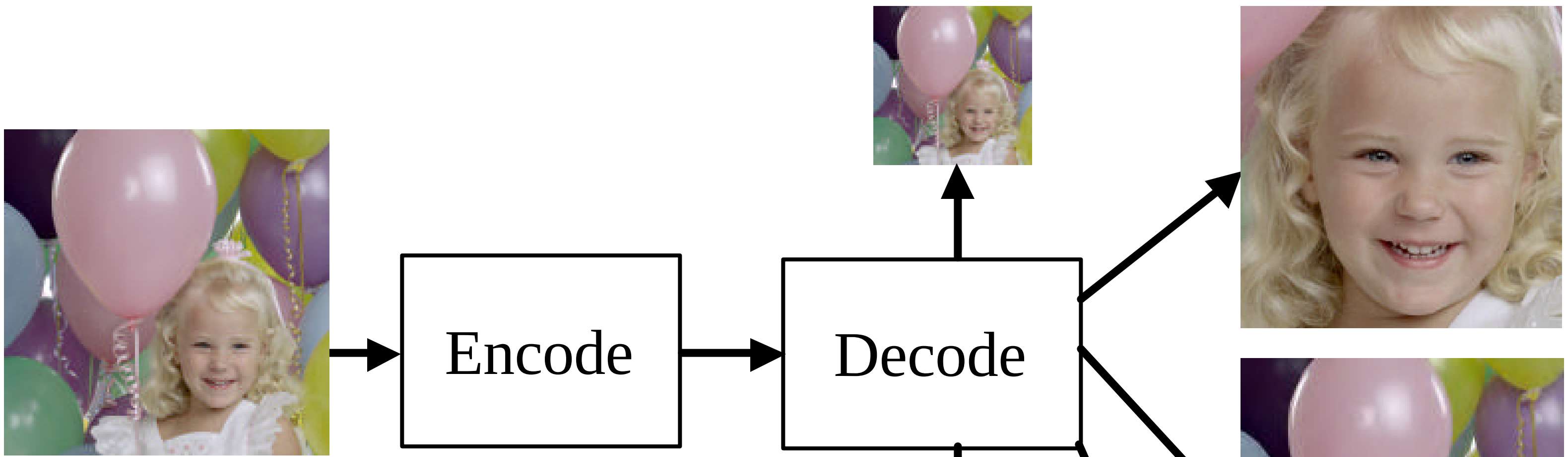
JPEG 2000 Timetable (Parts 1-7)

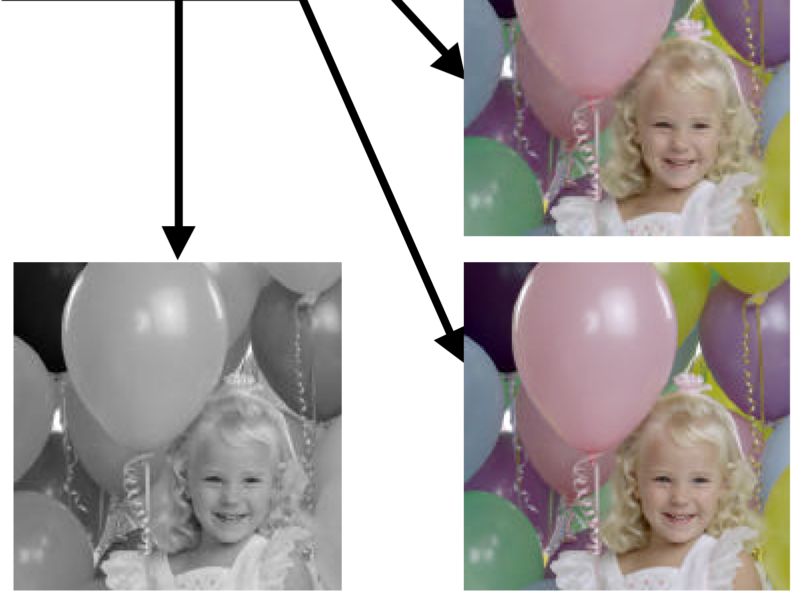
Part	Title	CFP	WD	CD	FCD	FDIS	IS
1	JPEG 2000 Image Coding System: Core Coding Sys	97/03	99/03	99/12	00/03	00/10	00/12
2	JPEG 2000 Image Coding System: Extensions	97/03	00/03	00/08	00/12	01/07	01/11
3	Motion JPEG 2000	99/12	00/07	00/12	01/03	01/07	01/11
4	Conformance Testing	99/12	00/07	00/12	01/07	02/02	02/05
5	Reference Software	99/12	00/03	00/07	00/12	01/08	01/11
6	Compound Image File Format	97/03	00/12	01/03	01/11	02/11	03/01
7	---Withdrawn---						

JPEG 2000 Timetable (Parts 8-12)

Part	Title	CFP	WD	CD	FCD	FDIS	IS
8	JPSEC: Secure JPEG2000	02/03	02/10	03/03	03/07	03/11	04/03
9	JPIP: Interactivity Tools, APIs and Protocols	02/03	02/07	03/03	03/07	03/11	04/03
10	JP3D: 3-D and Floating Point Data	02/03	02/10	03/03	03/07	03/11	04/03
11	JPWL: Wireless	02/07	03/03	03/07	03/11	04/02	04/05
12	ISO Base Media File Format	02/10	02/10	02/10	02/10	02/11	03/01

JPEG2000 Compression Paradigm





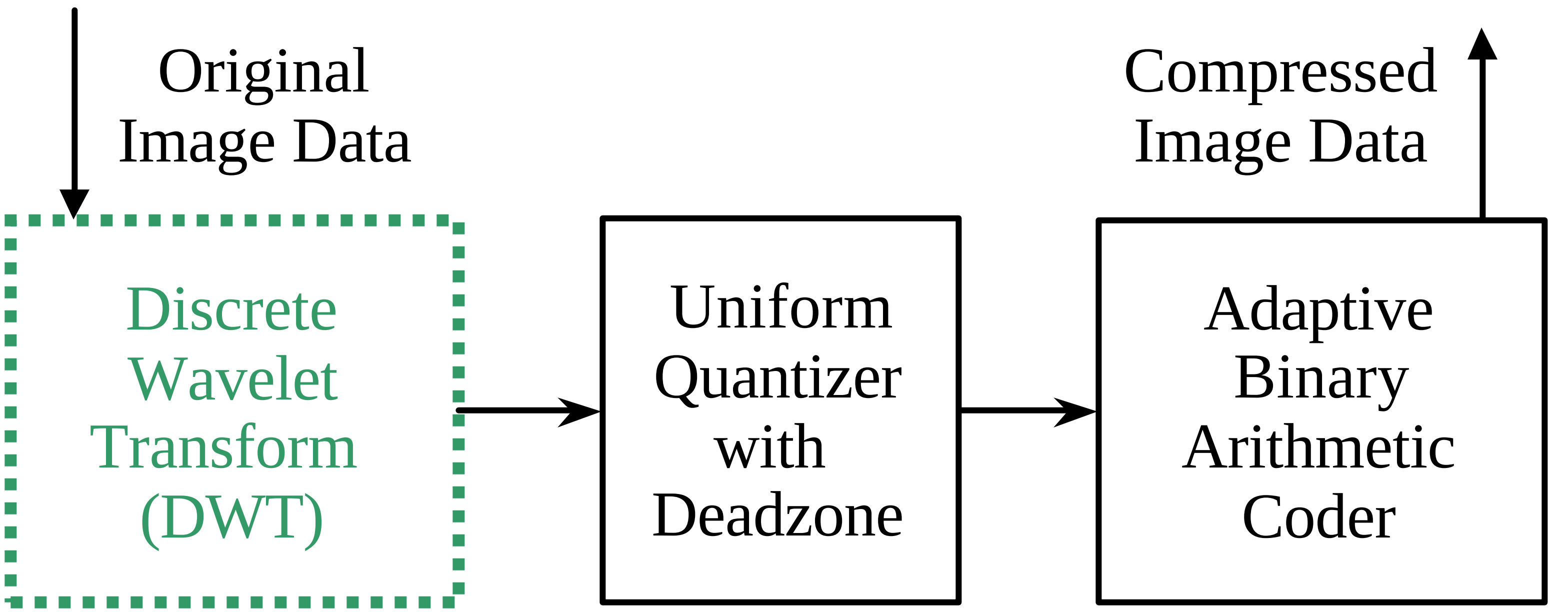
<u>Decode Choices</u>

Image resolution
SNR fidelity
Visual fidelity
Target filesize
Lossless/lossy
Region-of-interest
Tiles

JPEG2000 Features

- Improved compression efficiency (estimated 5-30% depending on the image size and bit rate)
- Lossy to lossless
- Multiple resolution
- Embedded bit stream (progressive decoding)
- Region of interest coding (ROI)
- Error resilience
- Bit stream syntax
- File format

JPEG2000 Part I Encoder



Benefits of DWT

- Multiple resolution representation
- Lossless representation with integer filters
- Better decorrelation than DCT, resulting in higher compression efficiency
- Use of visual models
  - DWT provides a frequency band decomposition of the image where each subband can be quantized according to its visual importance (similar to the quantization table specification in JPEG-DCT)

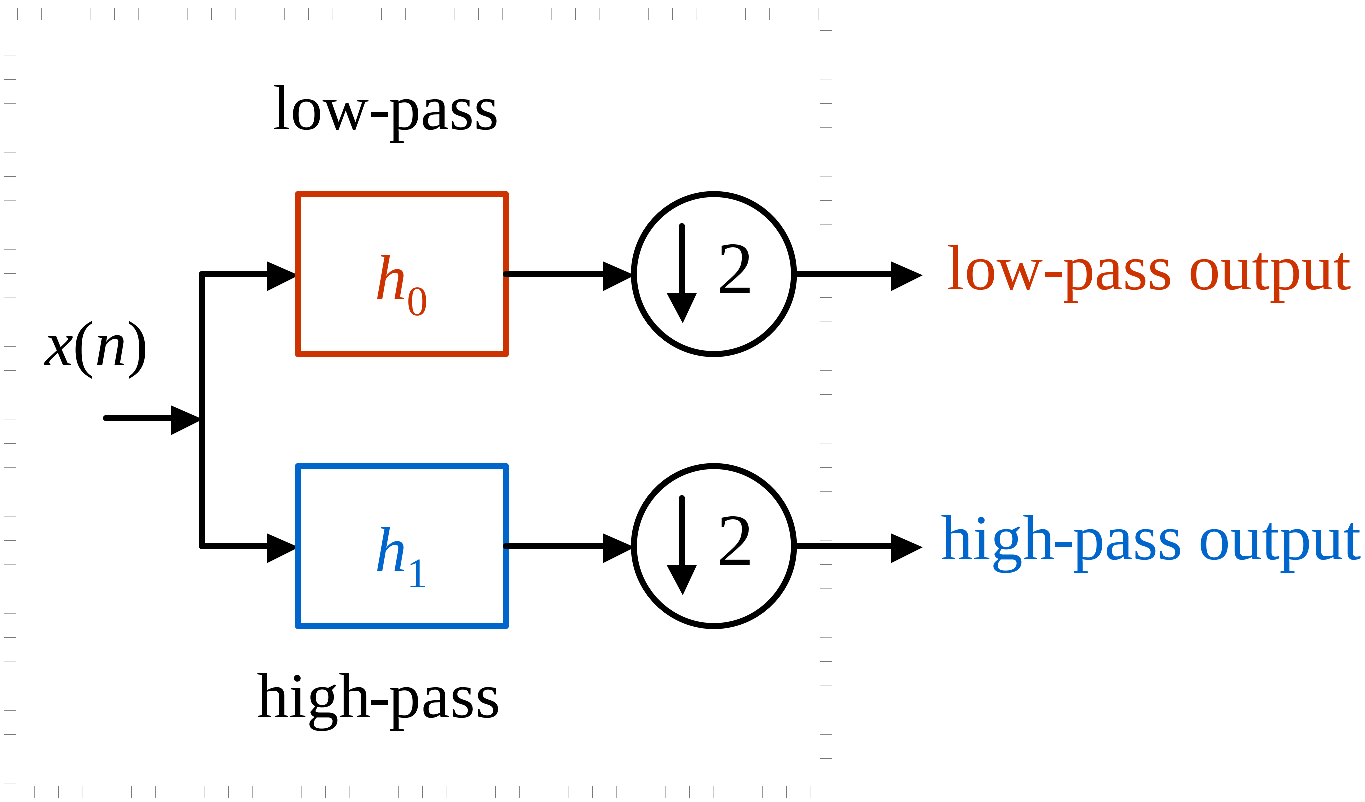
JPEG2000 DWT Choices

- JPEG-2000 Part I only allows successive powers of two splitting of the LL band and the use of two DWT filters:
  - The integer (5,3) filter provides fast implementation (faster than DCT) and lossless capability, but at the expense of some loss in coding efficiency.
  - The Daubechies (9,7) floating-point filter that provides superior coding efficiency. The analysis filters are normalized to a DC gain of one and a Nyquist gain of 2.
- Part II allows for arbitrary size filters (user-specified in the header), arbitrary wavelet decomposition trees, and different filters in the horizontal vs. vertical directions.

Computational Complexity

- 1-D decomposition:
  - (9,7) filter-bank: 3 mults + 4 adds per input sample.
  - (5,3) filter-bank: 1 mult. + 2 adds per input sample.
- 2-D decomposition:
  - (9,7) filter-bank: 6 mults + 8 adds per input sample.
  - (5,3) filter-bank: 2 mult. + 4 adds per input sample.
- For a 5 level octave decomposition, the complexity increases by a factor of 1.33
- JPEG: 2 mults + 6 adds per input sample.

The 1-D Two-Band DWT



Analysis filter bank

Example of Analysis Filter Bank

- 1-D signal:

  …100 100 100 100 200 200 200 200...

- Low-pass filter h_0: (-1 2 6 2 -1)/8

- High-pass filter h_1: (-1 2 -1)/2

- Before downsampling:

  … 100 100 87.5 112.5 187.5 212.5 200 200...

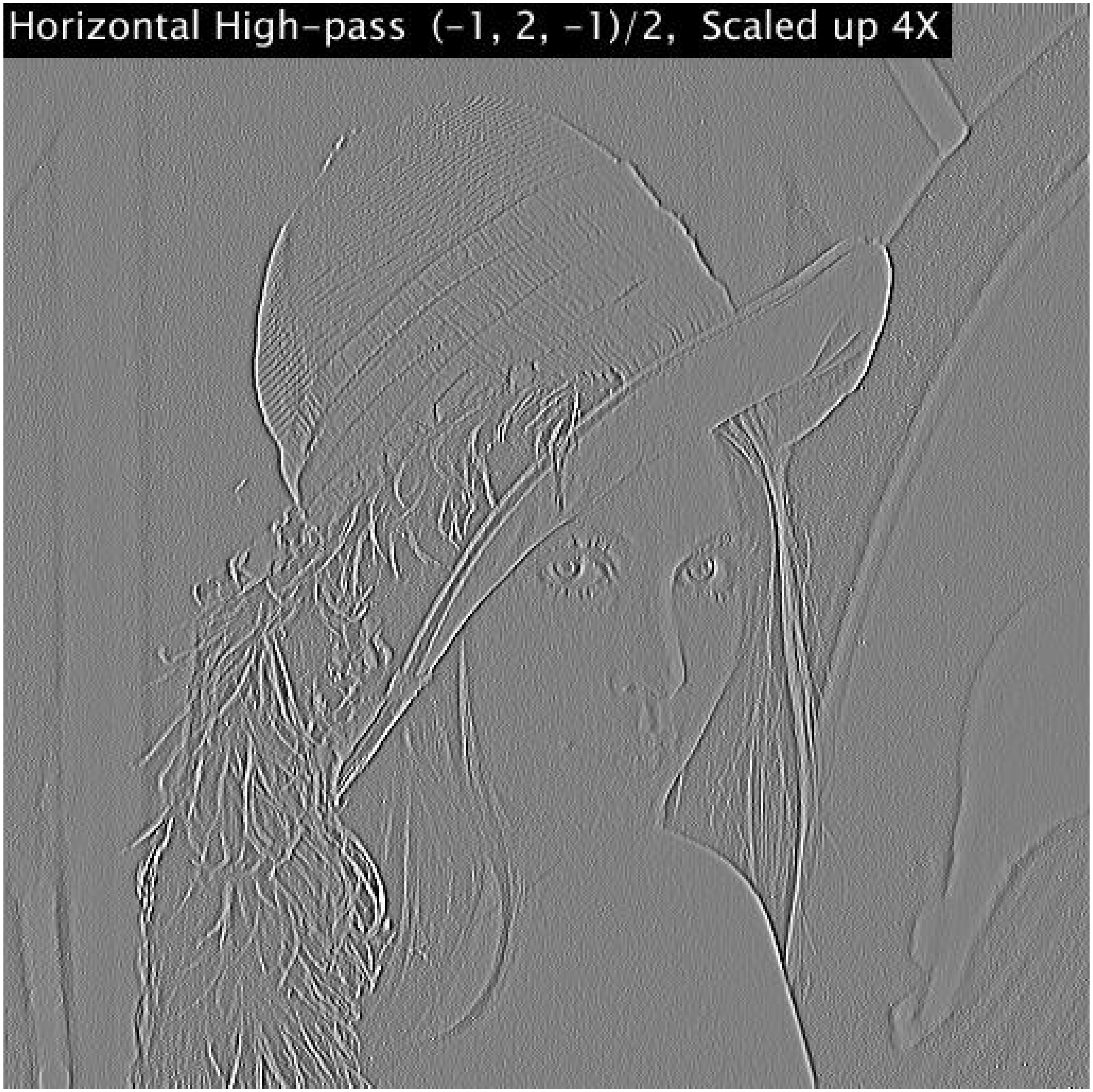
  … 0 0 0 -50 50 0 0 0…

- After downsampling:

  … 100 112.5 212.5 200...

  … 0 0 50 0 …


Horizontal Low-pass (-1, 2, 6, 2, -1)/8


Horizontal High-pass (-1, 2, -1)/2, Scaled up 4X

Inverse DWT

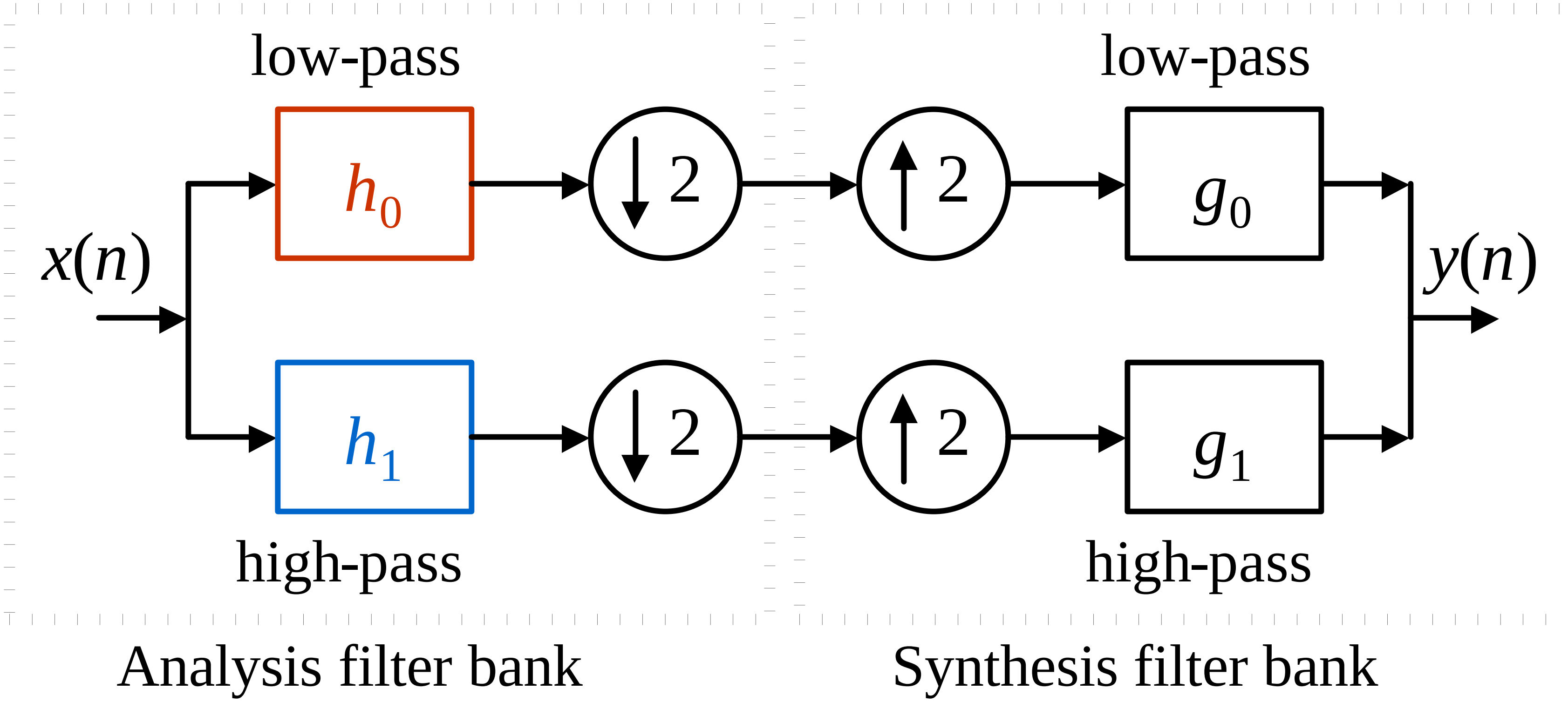
- During the inverse DWT, each subband is interpolated by a factor of two by inserting zeros between samples and then filtering each resulting sequence with the corresponding low-pass, g_0, or high-pass, g_1, **synthesis filters**.
- The filtered sequences are added together to form an approximation to the original signal.

… **0** 100 **0** 112.5 **0** 212.5 **0** 200…

… 0 **0** 0 **0** 50 **0** 0 **0**…

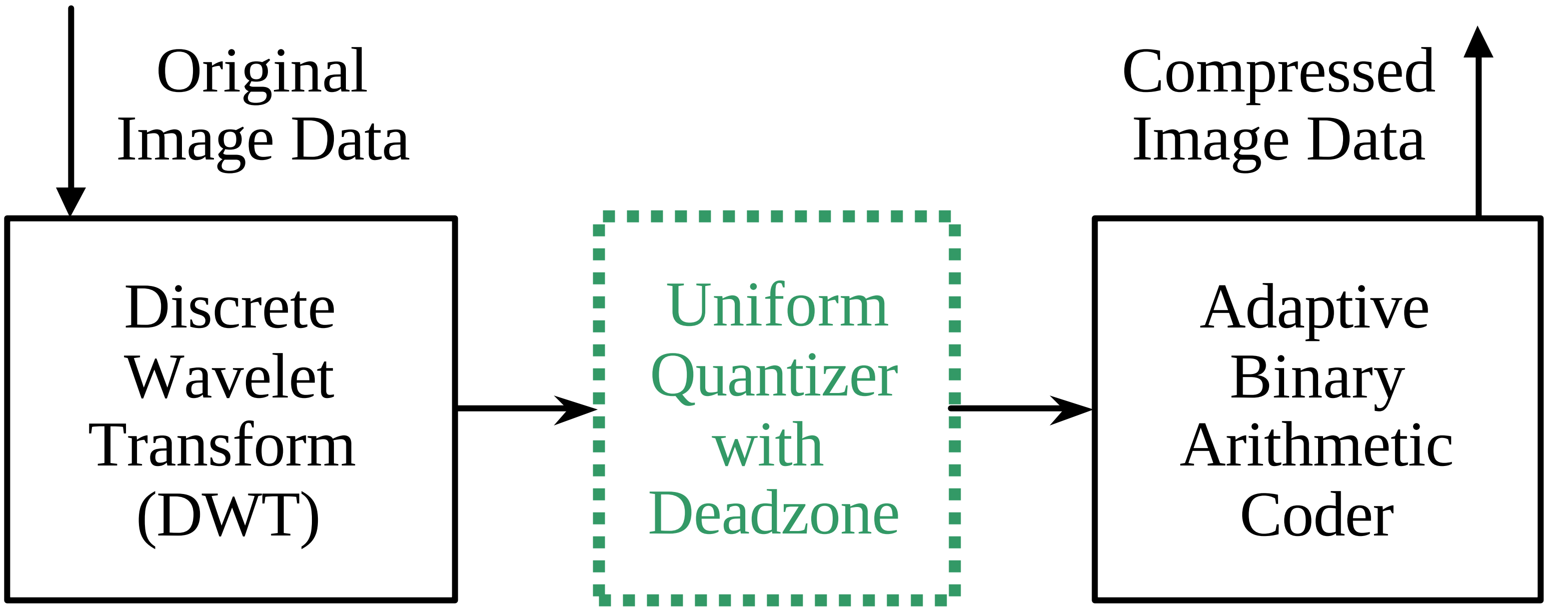
… 100 100 100 100 200 200 200 200...

The 1-D Two-Band DWT



Ideally, it is desired to choose the analysis filter banks (h_0 and h_1), and the synthesis filter banks (g_0 and g_1), in such a way so as to make the overall distortion zero, i.e., $x(n) = y(n)$. This is called the **perfect reconstruction** property.

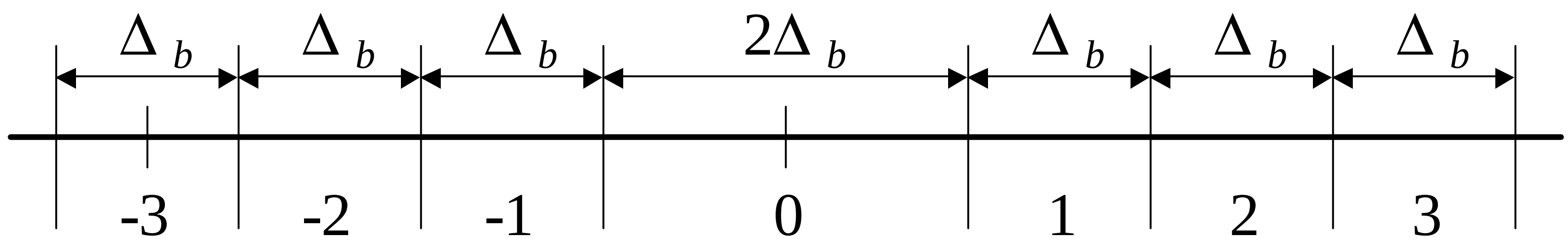
JPEG2000 Part I Encoder



Quantization step size can vary from one subband to another according to visual models (similar to JPEG Q-table specification).

Quantization in Part I

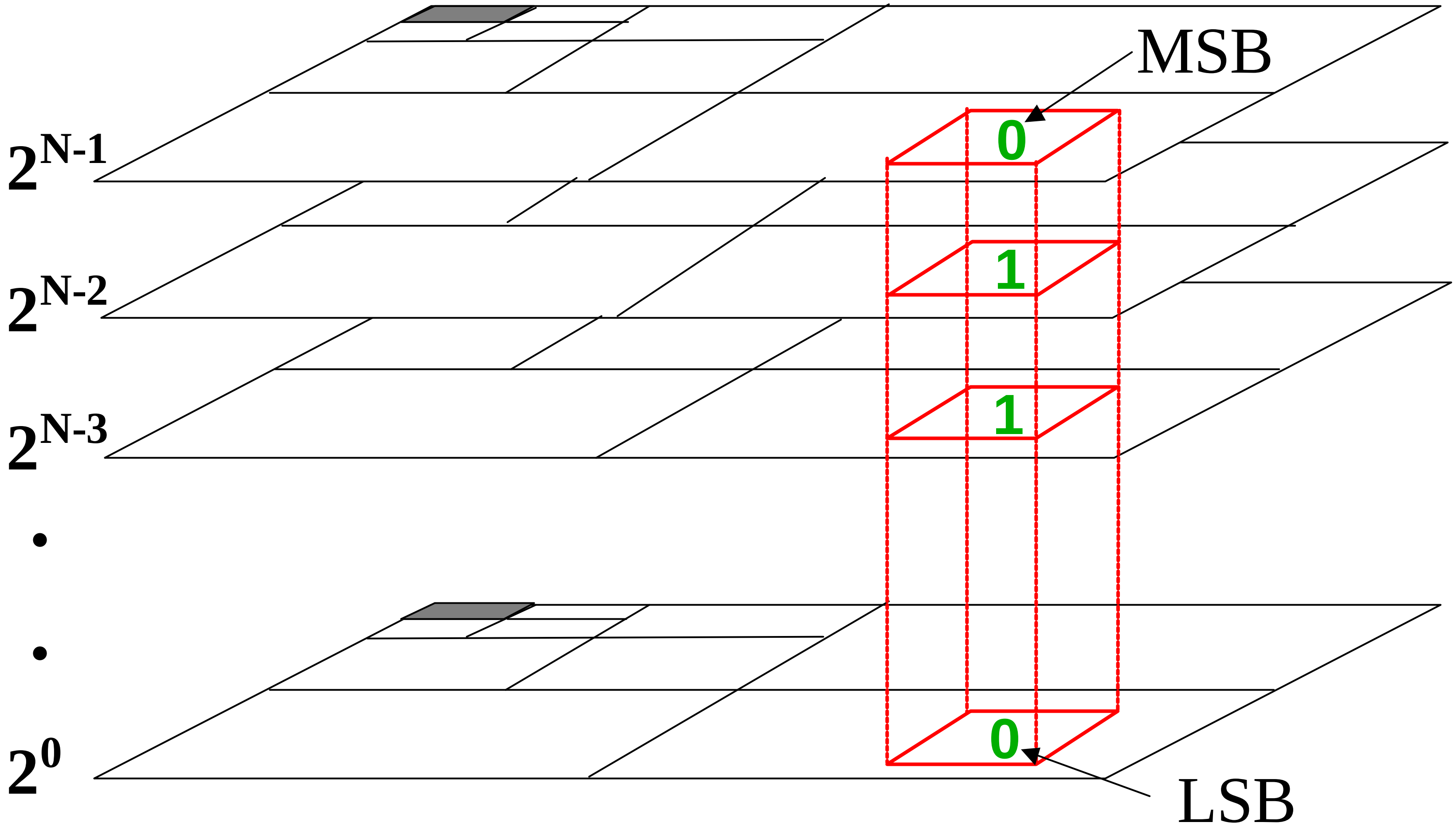
- Uniform quantization with deadzone is used to quantize all the wavelet coefficients.
- For each subband *b*, a basic quantizer step size Δ_b is selected by the user and is used to quantize all the coefficients in that subband.
- The choice of the quantizer step size for each subband can be based on visual models and is likened to the q-table specification in the JPEG DCT.



Embedded Quantization in Part 1

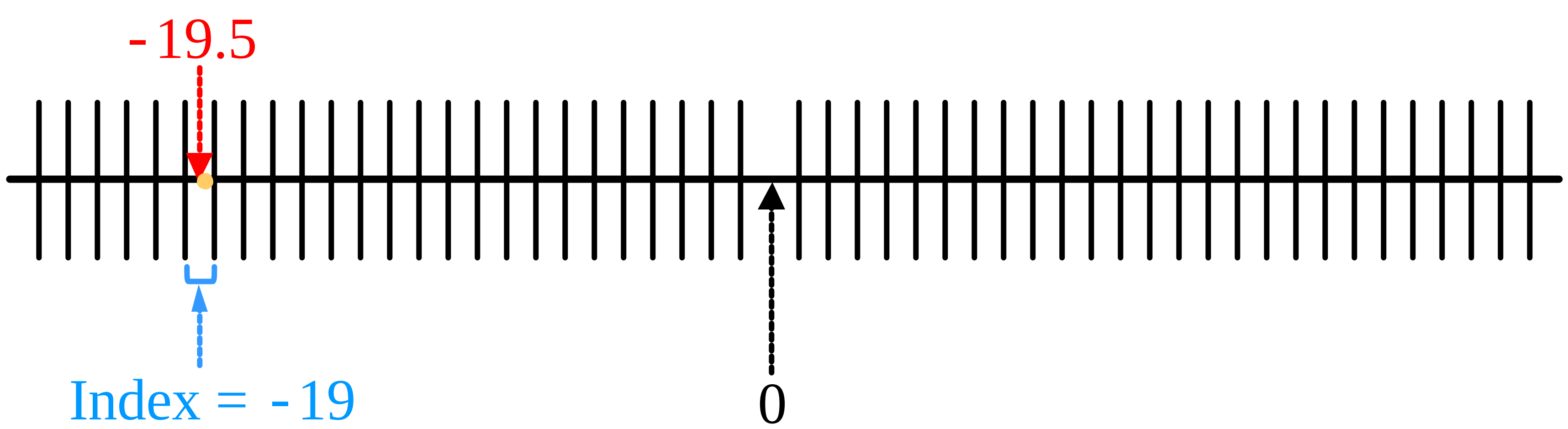
- Unlike JPEG Baseline, where the resulting quantizer index q is encoded as a single symbol, in JPEG2000 it is encoded one bit at a time, starting from the MSB and proceeding to the LSB.
- During this progressive encoding, the quantized wavelet coefficient is called **insignificant** if the quantizer index q is still zero. Once the first nonzero bit is encoded, the coefficient becomes **significant** and its sign is encoded.
- If the p least significant bits of the quantizer index still remain to be encoded, the reconstructed sample at that stage is identical to the one obtained by using a UTQ with deadzone with a step size of $\Delta_b 2^p$.

Embedded Quantization by Bit-Plane Coding



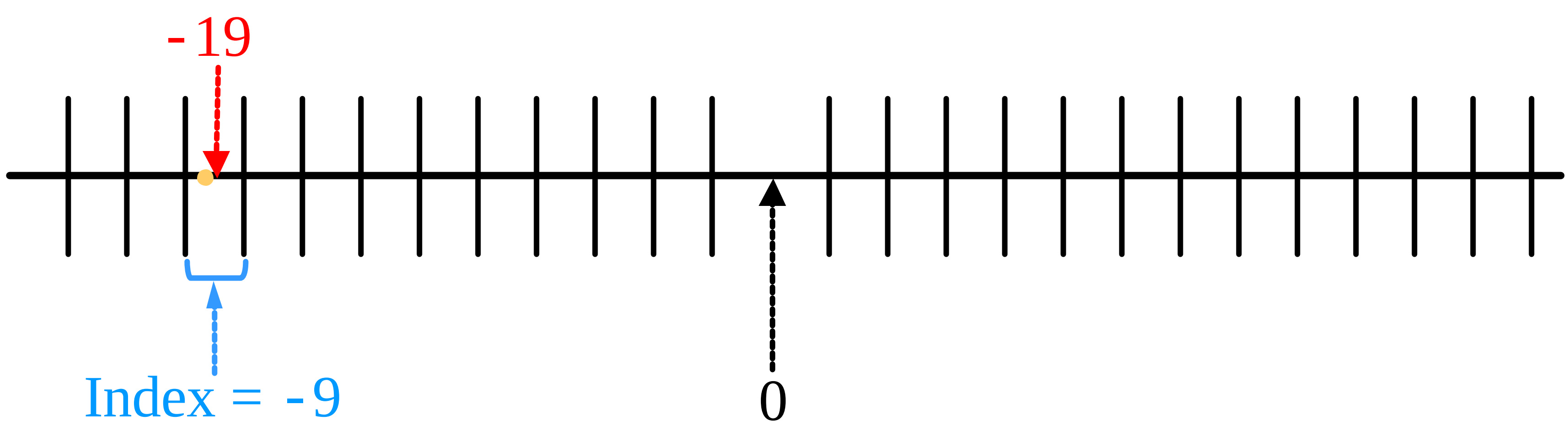
Embedded Quantization

- Wavelet coefficient value = -19.38
- Quantization stepsize = 1
- Quantized value = -19 = -00010011 (8-bit binary)
- Reconstructed value with full 8-bit index = -19.50



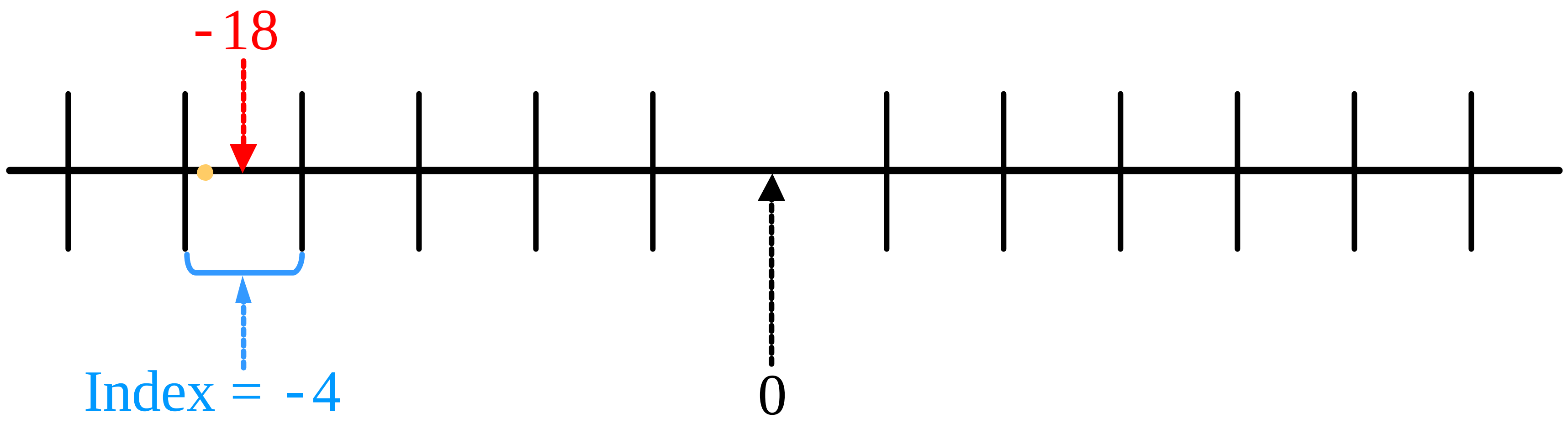
Embedded Quantization

- Original wavelet coefficient value = -19.38
- Quantization stepsize = 2
- Quantizer index value = -0001001? = -9
- Reconstructed value with partial 7-bit index = -19



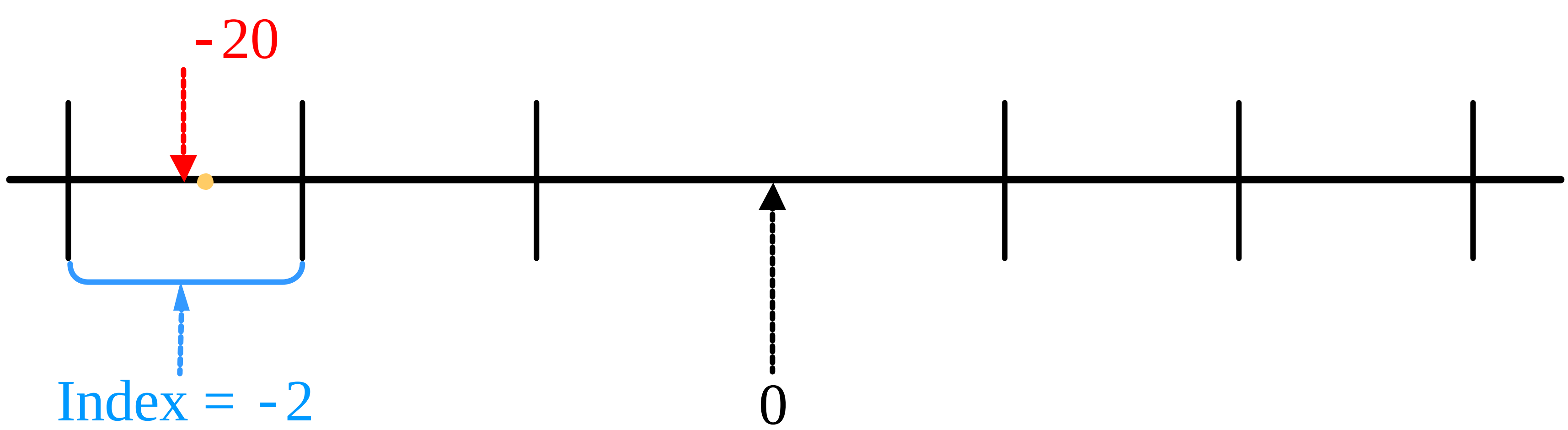
Embedded Quantization

- Original wavelet coefficient value = -19.38
- Quantization stepsize = 4
- Quantizer index value = -000100?? = -4
- Reconstructed value with partial 6-bit index = -18



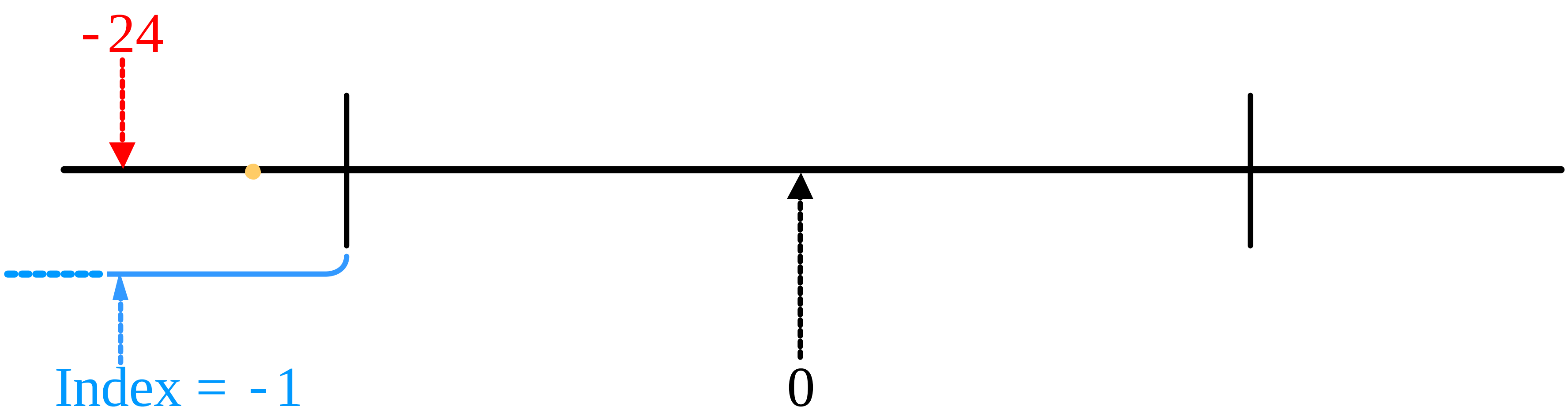
Embedded Quantization

- Original wavelet coefficient value = -19.38
- Quantization stepsize = 8
- Quantizer index value = -00010??? = -2
- Reconstructed value with partial 5-bit index = -20



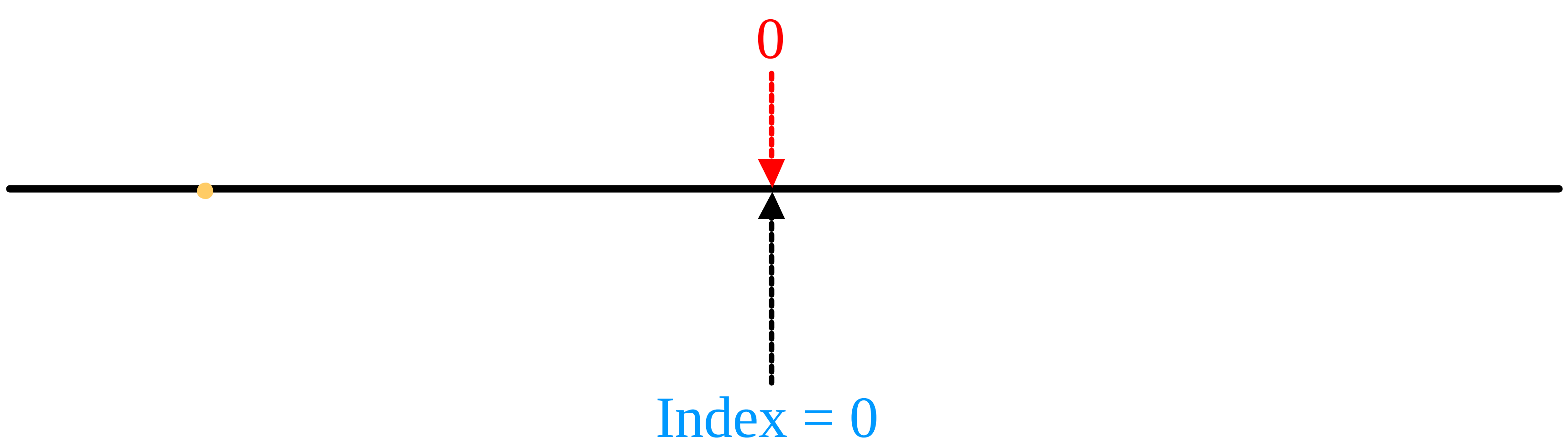
Embedded Quantization

- Original wavelet coefficient value = -19.38
- Quantization stepsize = 16
- Quantizer index value = -0001???? = -1
- Reconstructed value with partial 4-bit index = -24

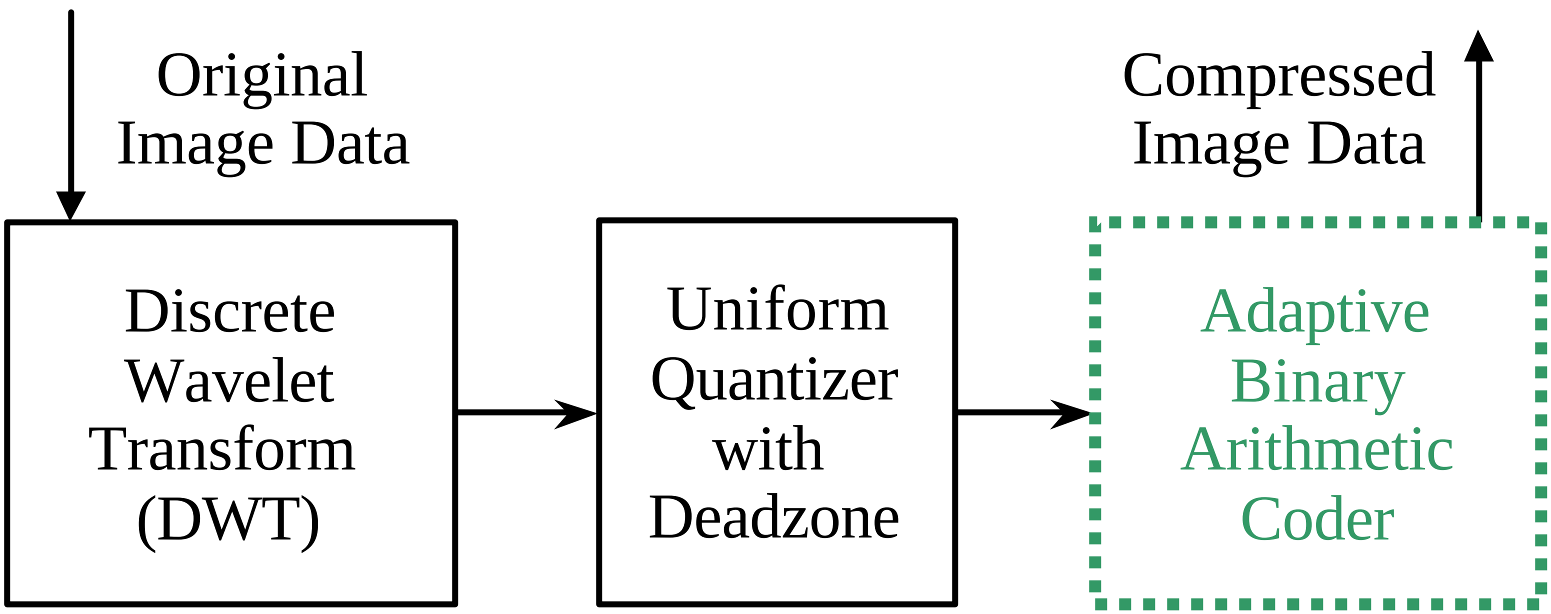


Embedded Quantization

- Original wavelet coefficient value = -19.38
- Quantization stepsize = 32
- Quantizer index value = 000????? = 0
- Reconstructed value with partial 3-bit index = 0



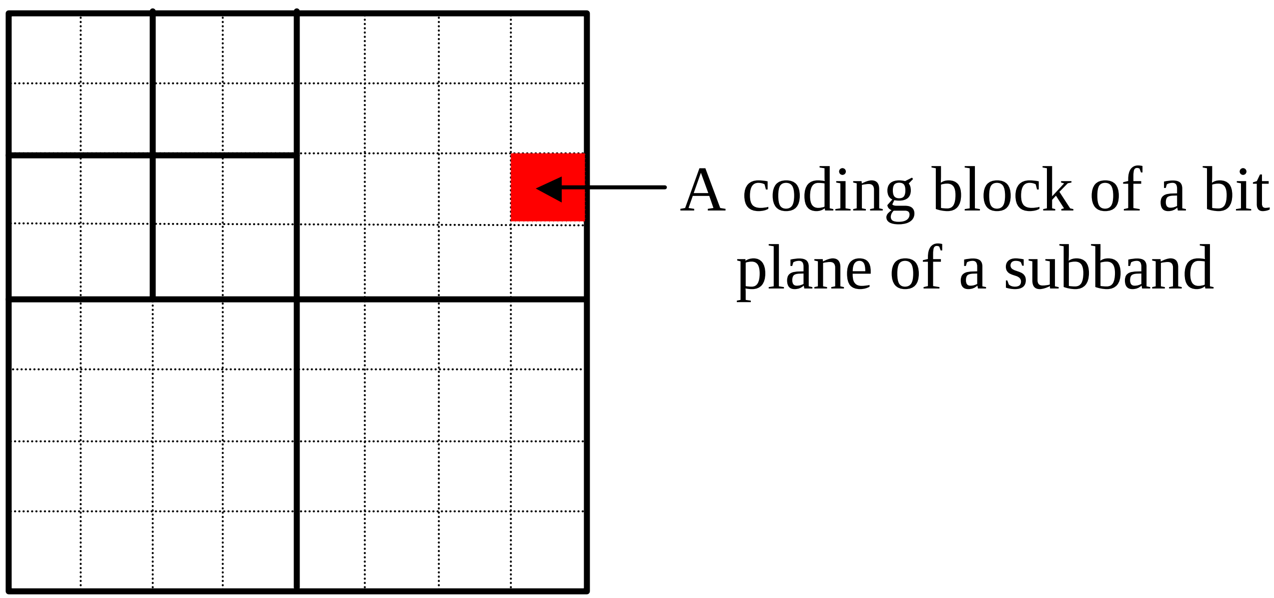
JPEG2000 Part I Encoder



Context-based adaptive binary arithmetic coding is used in JPEG2000 to efficiently compress each individual bit plane.

JPEG2000 Entropy Coder

- Each bit plane is further broken down into blocks (e.g., 64 x 64). The blocks are coded independently (i.e., the bit stream for each block can be decoded independent of other data) using three coding passes. The coding progresses from the most significant bit-plane to the least significant bit-plane.



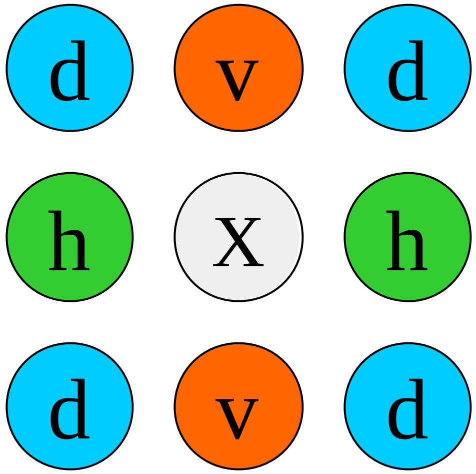
JPEG2000 Entropy Coder

- The binary value of a sample in a block of a bit plane of a subband is coded as a binary symbol with the JBIG2 **MQ-Coder** that is a context-based adaptive arithmetic coder.
- Each bit-plane of each block of a subband is encoded in **three sub bit plane passes** instead of a single pass. The bitstream can be truncated at the end of each pass. This allows for:
  - Optimal embedding, so that the information that results in the most reduction in distortion for the least increase in file size is encoded first.
  - A larger number of bit-stream truncation points to achieve finer SNR scalability.

Significance Propagation Pass

- The first pass in a new bit plane is called the **significance propagation pass**. A symbol is encoded if it is insignificant but at least one of its eight-connected neighbors is significant as determined from the previous bit plane and the current bit plane based on coded information up to that point. These locations have the highest probability of becoming significant).

- The probability of the binary value at a given location of a bit-plane of a block of a subband is modeled by a context formed from the significance values of its neighbors.



Refinement and Clean-up Passes

- **Refinement** (REF): Next, the significant coefficients are refined by their bit representation in the current bit-plane.
- **Clean-up**: Finally, all the remaining coefficients in the bit-plane are encoded. (*Note*: the first pass of the MSB bit-plane of a subband is always a clean-up pass).
- The coding for the first and third passes are identical, except for the run coding that is employed in the third pass.
- The maximum number of contexts used in any pass is no more than nine, thus allowing for extremely rapid probability adaptation that decreases the cost of independently coded segments.

Sig. Prop.	=	0
Refine	=	0
Cleanup	=	21
Total Bytes		21

Bit plane **1**

Compression ratio = 12483 : 1

RMSE = 39.69 PSNR = 16.16 db

% refined = 0 % insig. = 99.99

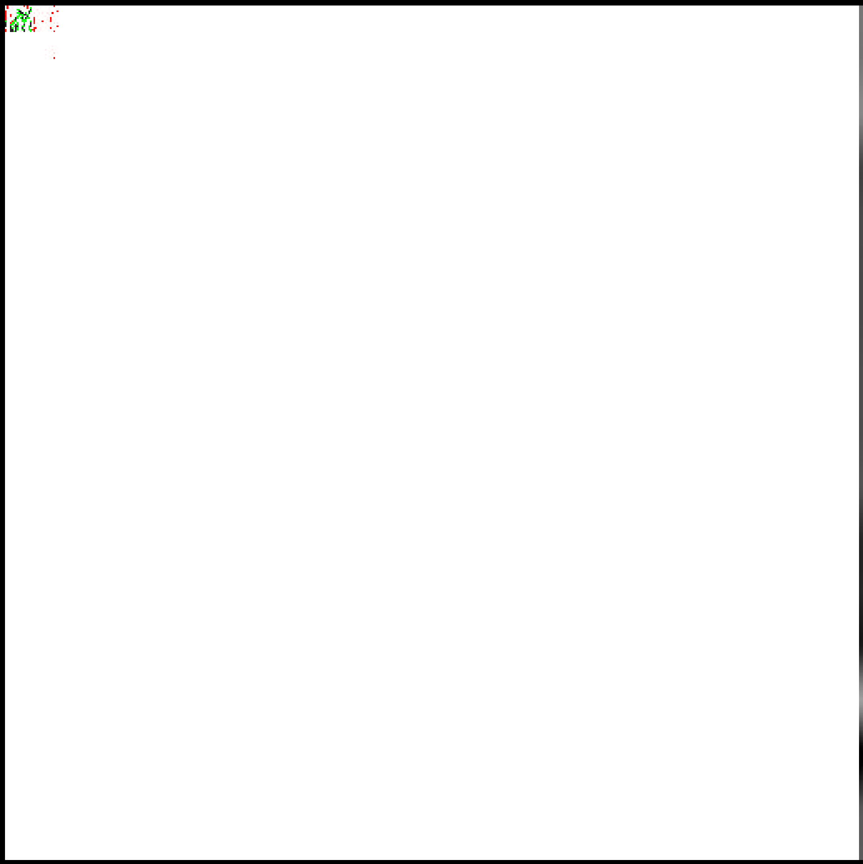
Sig. Prop.	=	18
Refine	=	0
Cleanup	=	24
Total Bytes		42

Bit plane **2**

Compression ratio = 4161 : 1

RMSE = 29.11 PSNR = 18.85 db

% refined = 0.01 % insig. = 99.95



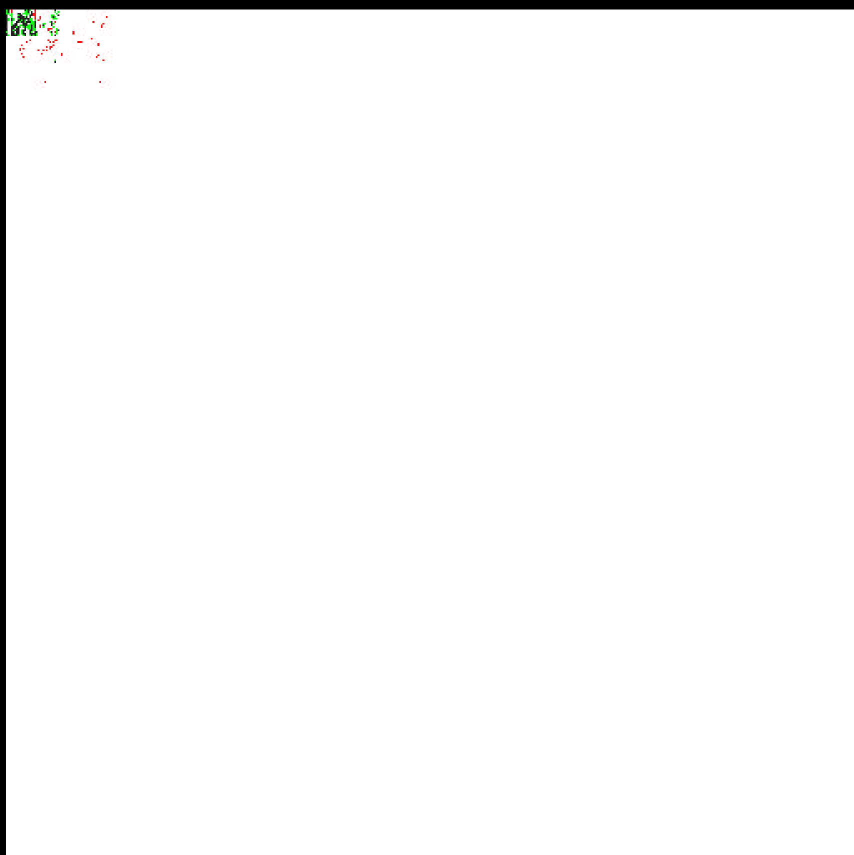
Sig. Prop.	=	38
Refine	=	13
Cleanup	=	57
Total Bytes		108

Bit plane **3**

Compression ratio = 1533 : 1

RMSE = 21.59 PSNR = 21.45 db

% refined = 0.05 % insig. = 99.89



Sig. Prop.	=	78	Bit plane	**4**
Refine	=	37	Compression ratio = 593 : 1	
Cleanup	=	156	RMSE = 16.58	PSNR = 23.74 db
Total Bytes		271	% refined = 0.11	% insig. = 99.77

Sig. Prop.	=	224
Refine	=	73
Cleanup	=	383
Total Bytes		680

Bit plane **5**

Compression ratio = 233 : 1

RMSE = 12.11 PSNR = 26.47 db

% refined = 0.23 % insig. = 99.43

Sig. Prop. = 551

Refine = 180

Cleanup = 748

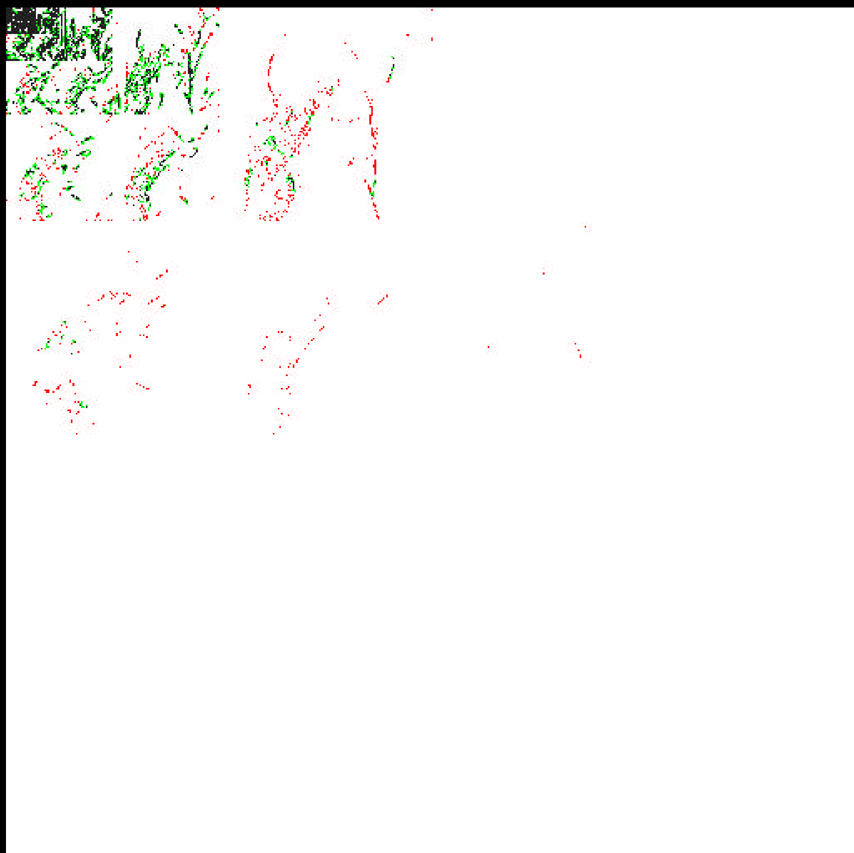
Total Bytes 1479

Bit plane **6**

Compression ratio = 101 : 1

RMSE = 8.65 PSNR = 29.39 db

% refined = 0.58 % insig. = 98.68



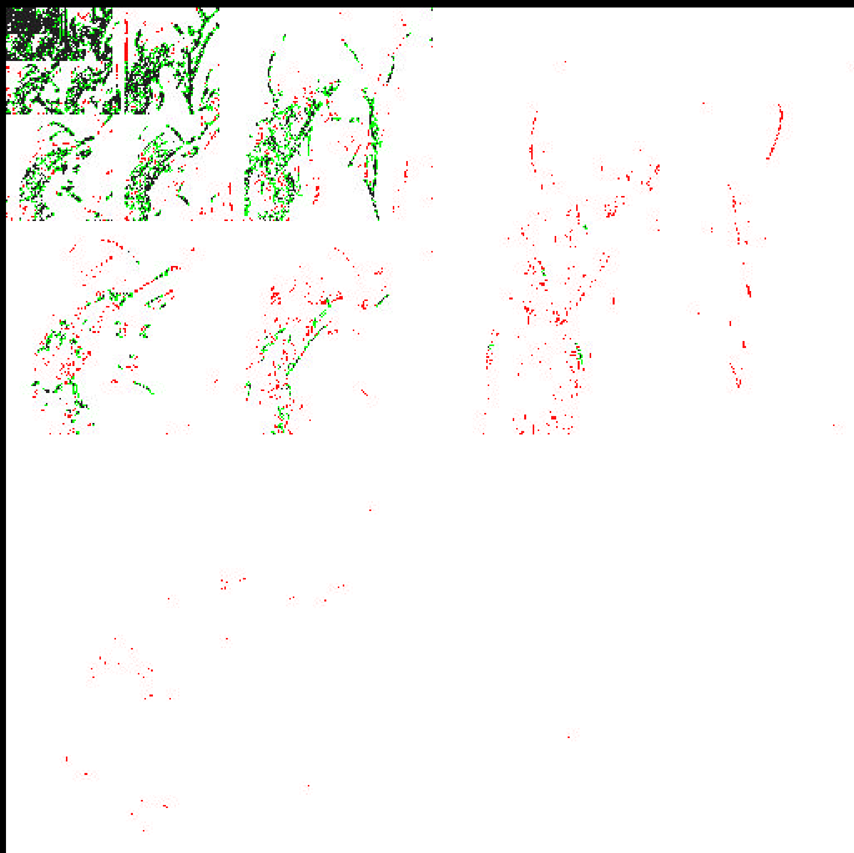
Sig. Prop.	=	1243
Refine	=	418
Cleanup	=	1349
Total Bytes		3010

Bit plane **7**

Compression ratio = 47 : 1

RMSE = 6.02 PSNR = 32.54 db

% refined = 1.32 % insig. = 97.09



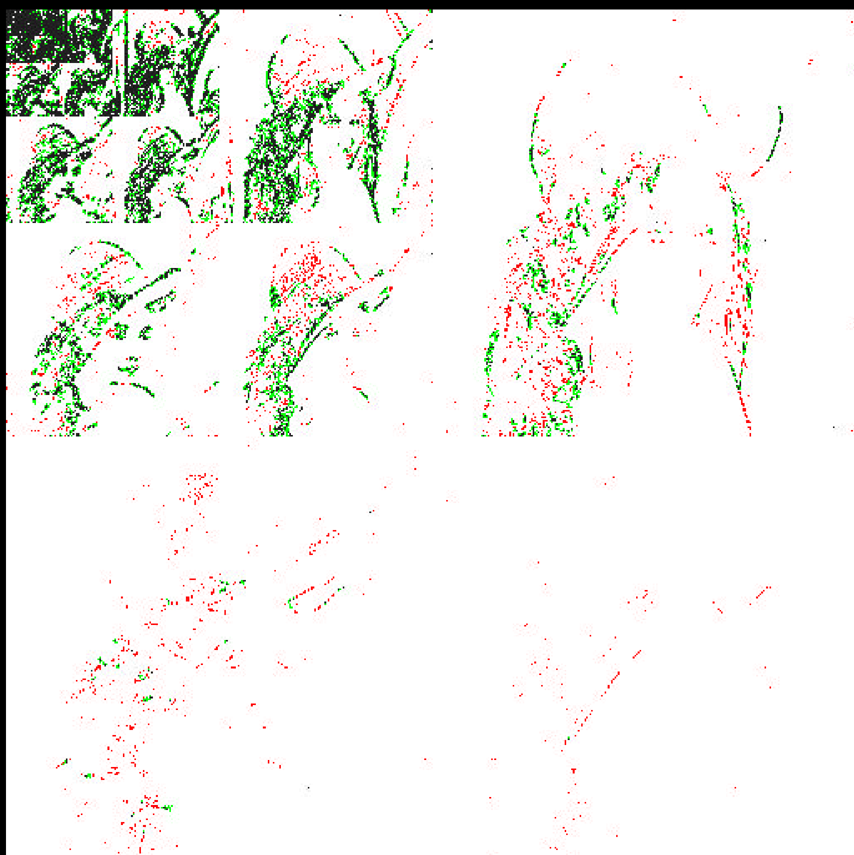
Sig. Prop.	=	2315
Refine	=	932
Cleanup	=	2570
Total Bytes		5817

Bit plane **8**

Compression ratio = 23 : 1

RMSE = 4.18   PSNR = 35.70 db

% refined = 2.91   % insig. = 93.99



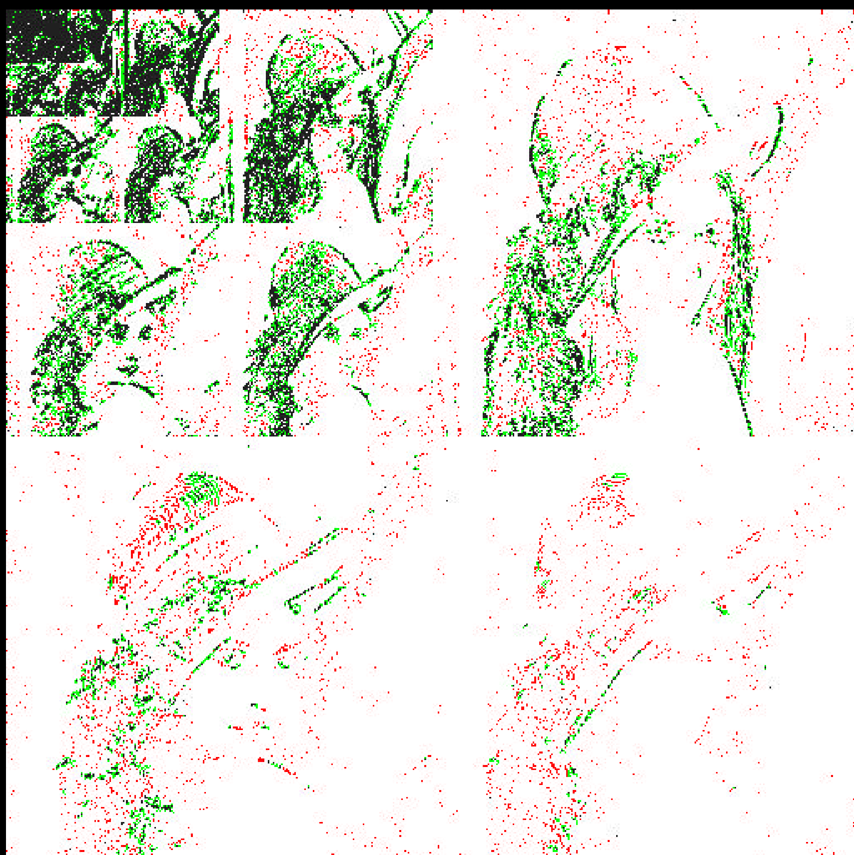
Sig. Prop.	=	4593
Refine	=	1925
Cleanup	=	5465
Total Bytes		11983

Bit plane **9**

Compression ratio = 11.2 : 1

RMSE = 2.90 PSNR = 38.87 db

% refined = 6.01 % insig. = 87.66



Sig. Prop. = 10720
Refine = 3917
Cleanup = 12779

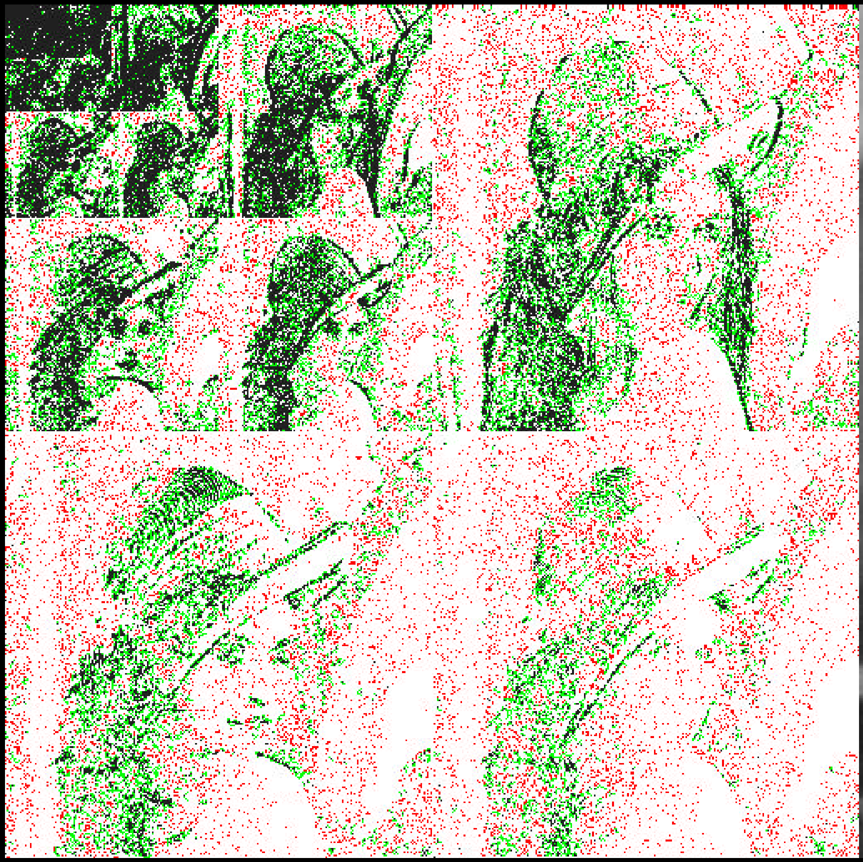
Total Bytes 27416

Bit plane **10**

Compression ratio = 5.16 : 1

RMSE = 1.78 PSNR = 43.12 db

% refined = 12.34 % insig. = 82.00



Sig. Prop. = 25421
Refine = 8808
Cleanup = 5438

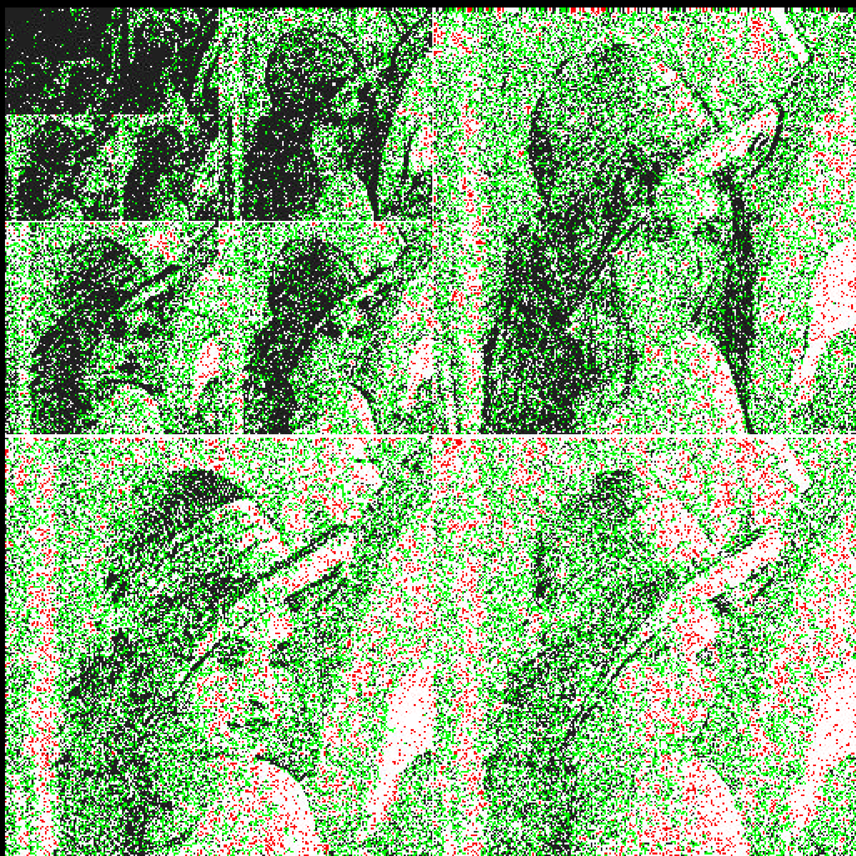
Total Bytes 39667

Bit plane **11**

Compression ratio = 2.90 : 1

RMSE = 0.90 PSNR = 49.00 db

% refined = 28.12 % insig. = 46.80



Wavelet Bit Plane Compression Table

BP	JPEG-2K Bytes†	JPEG-2K PSNR-dB	JPEG-2K RMSE	JPEG Bytes‡	JPEG PSNR	JPEG RMSE
1	21	16.16	39.69			
2	63	18.85	29.11			
3	171	21.45	21.59			
4	442	23.74	16.58			
5	1122	26.47	12.11			
6	2601	29.39	8.65			
7	5821*	32.54	6.02	5804*	29.77	8.28
8	11680*	35.70	4.18	11696*	33.33	5.49
9	23716*	38.87	2.90	23932*	36.50	3.81
10	51164*	43.12	1.78	51636*	40.16	2.50
11	90855*	49.00	0.90	91216*	44.12	1.59

†6-level decomposition with (9,7) filter

‡IJG code with default q-table

*Filesize includes the header

Comparison of JPEG and JPEG-2000

JPEG, PSNR = 29.77 db
Filesize = 5804 Bytes

JPEG-2K, PSNR = 32.54 db
Filesize = 5821 Bytes (BP 7)

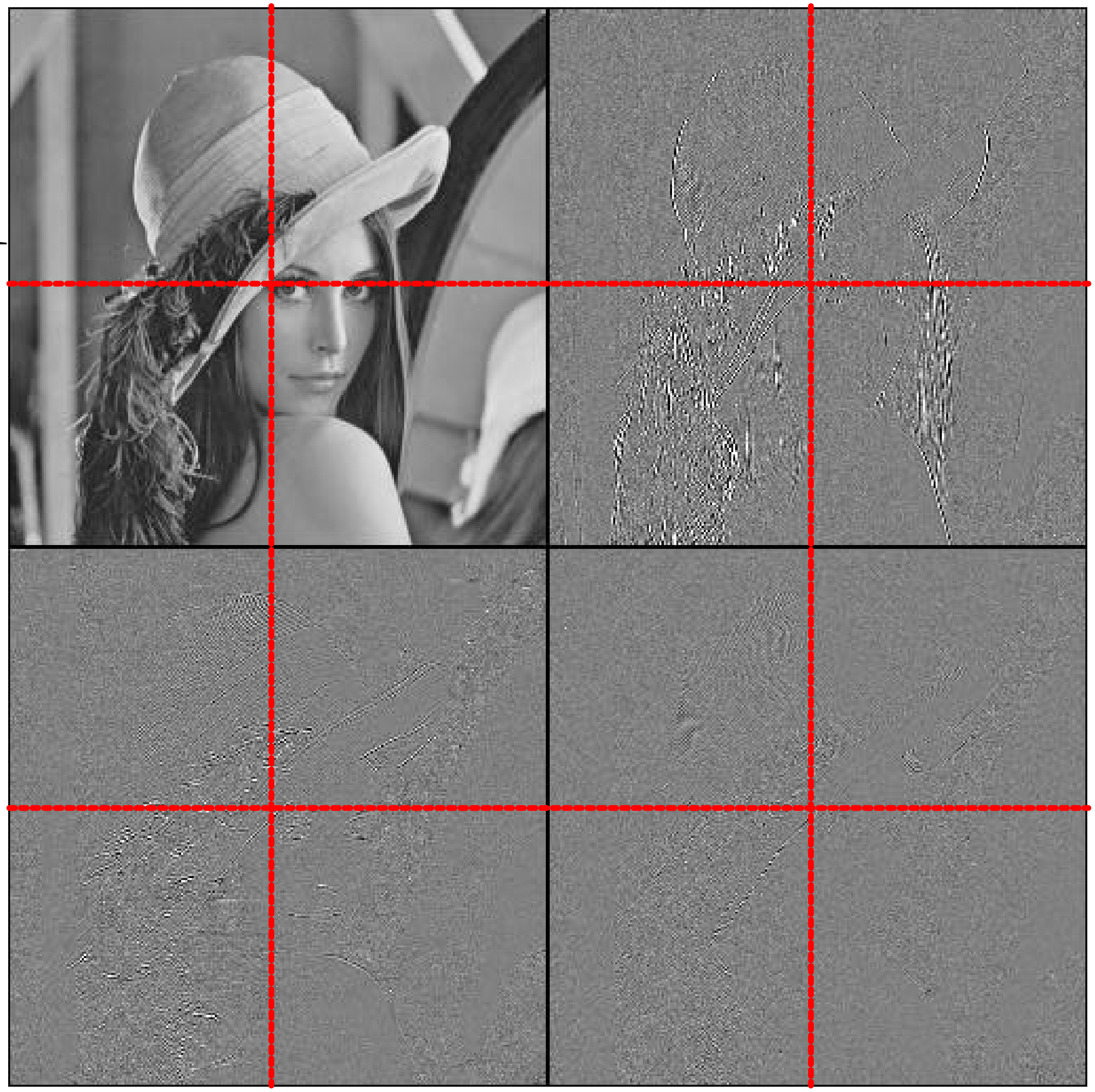
Benefits of the JPEG2000 Entropy Coder

- Bit plane representation of data allows for embedded bit stream
- Region of interest coding (ROI) coding can be performed by prioritizing the coding of the ROI bit plane information
- Arithmetic coding allows for efficient compression of sparse binary data
- Context modeling allows for efficient compression of binary correlated data
- Packetized information allows for improved error resilience

Example of Bit Plane Reordering

Subband 1

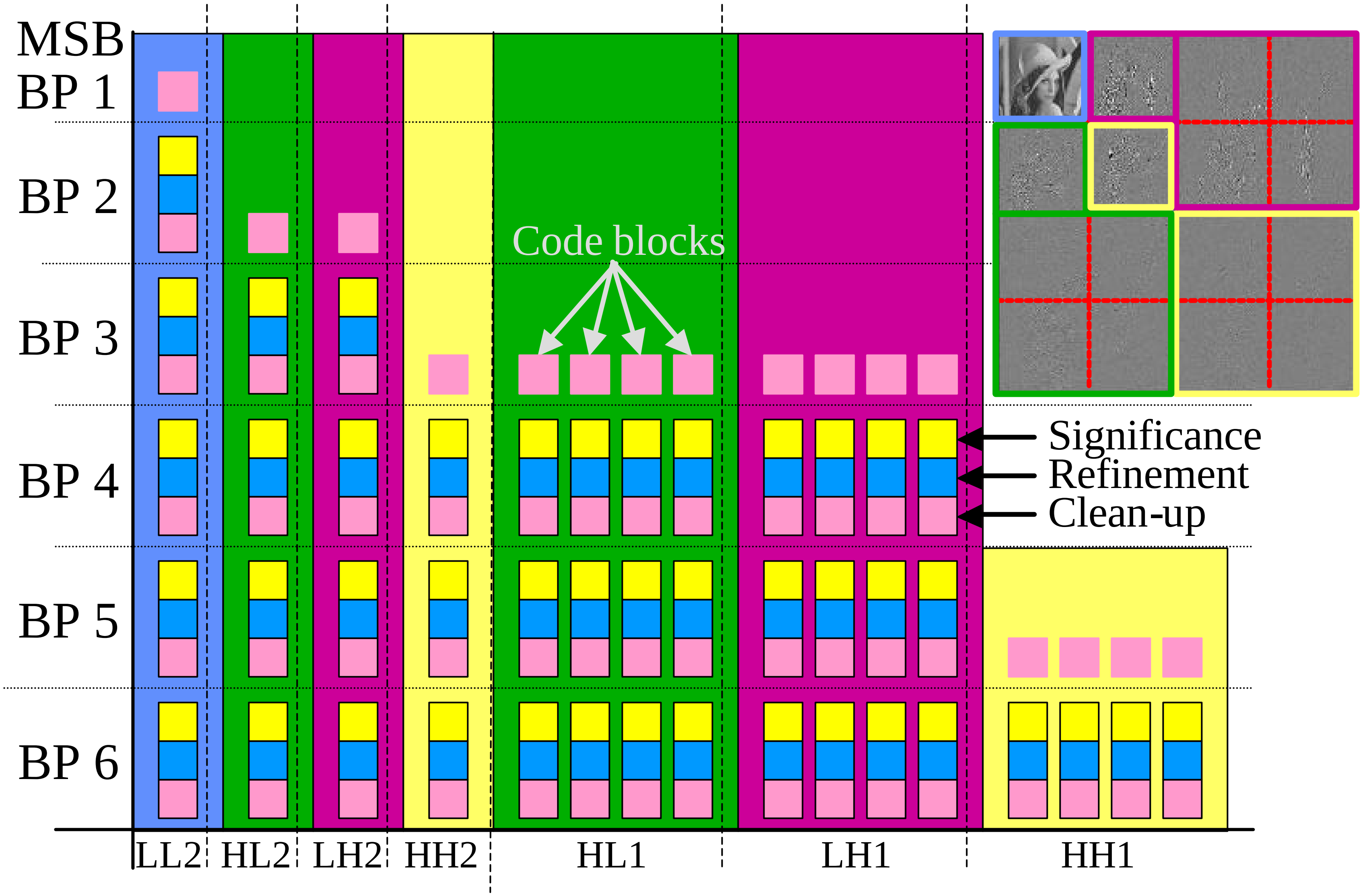
Subband 2



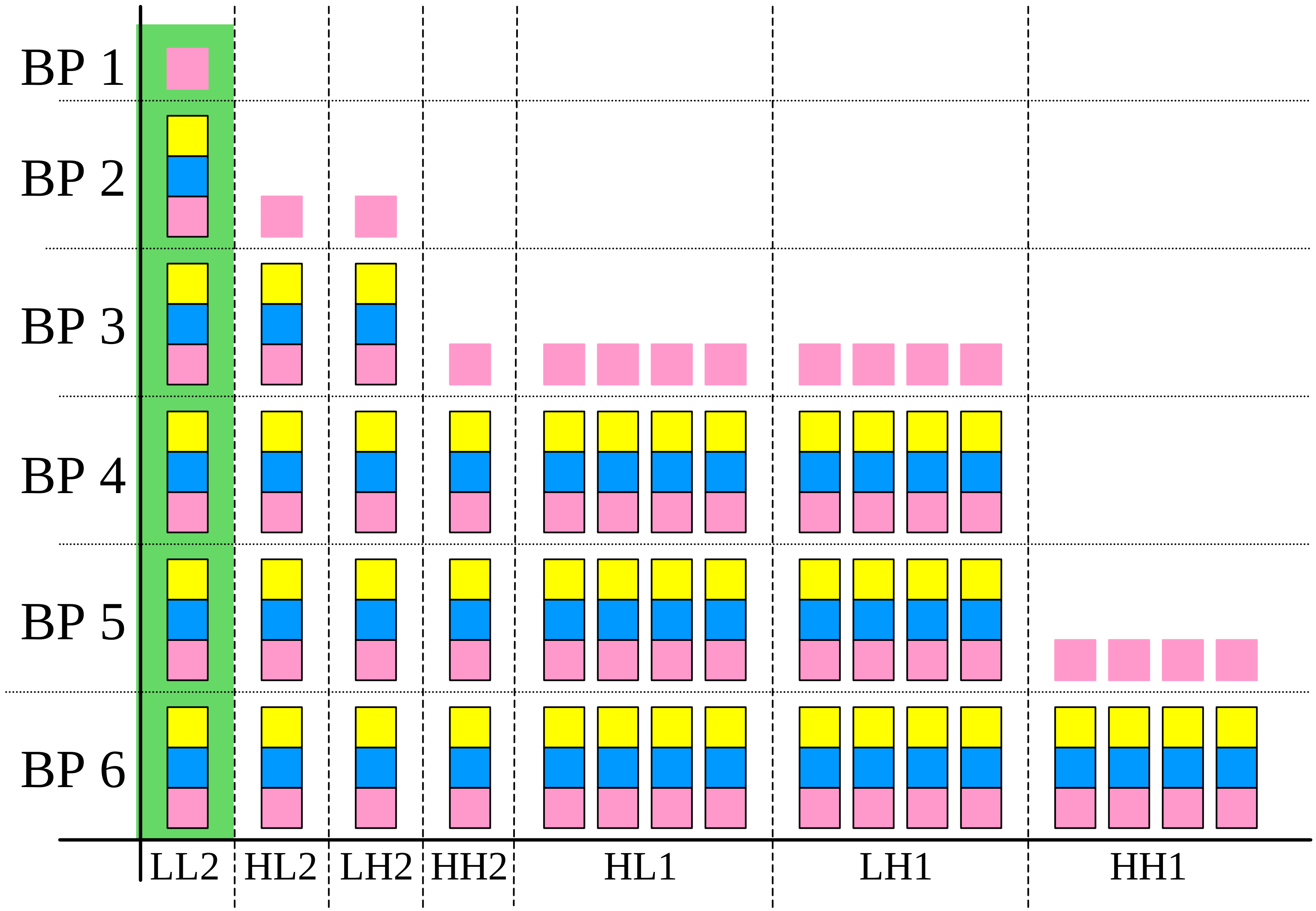
Subband 3

Subband 4

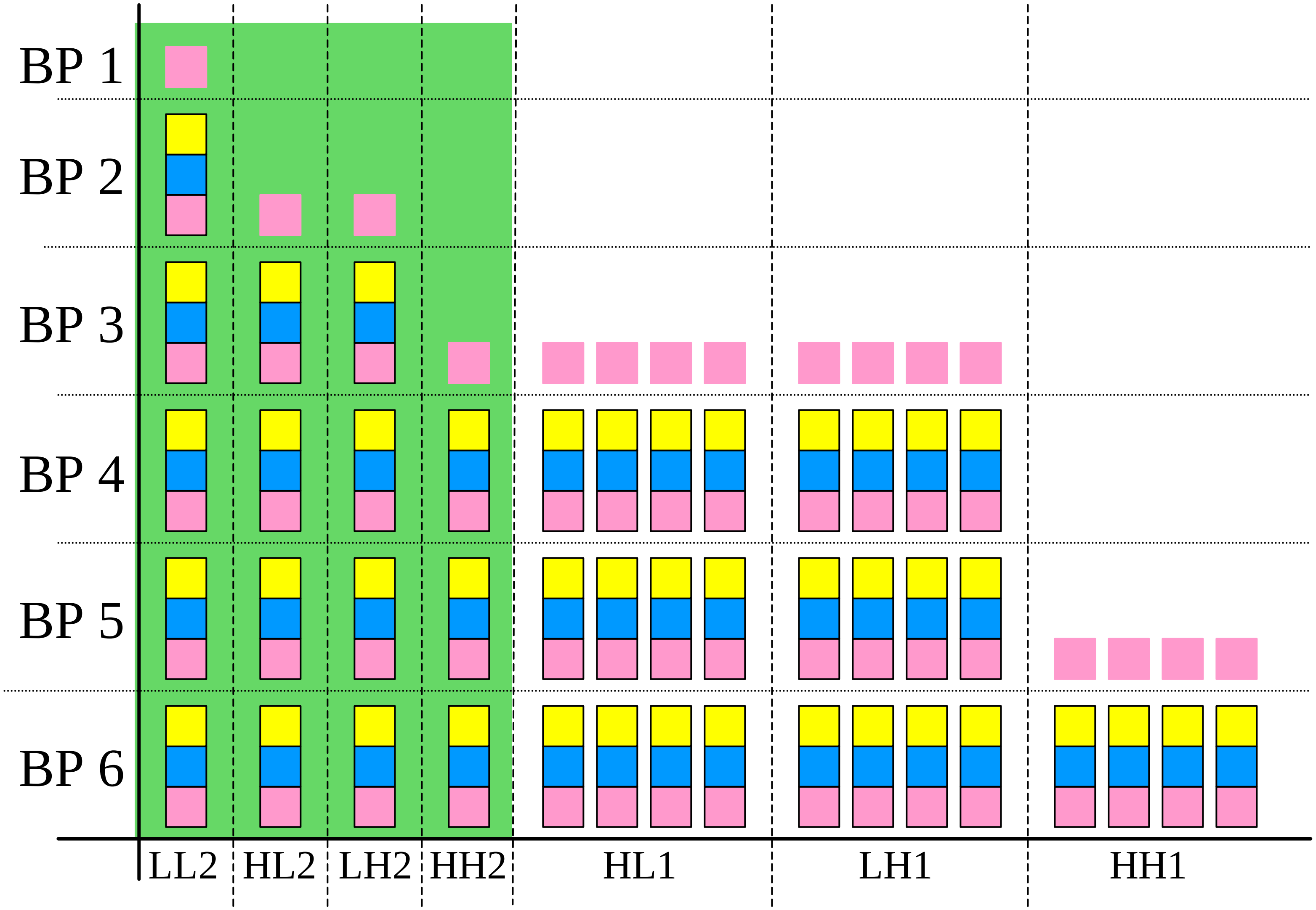
Example of Bit-Plane Data Ordering



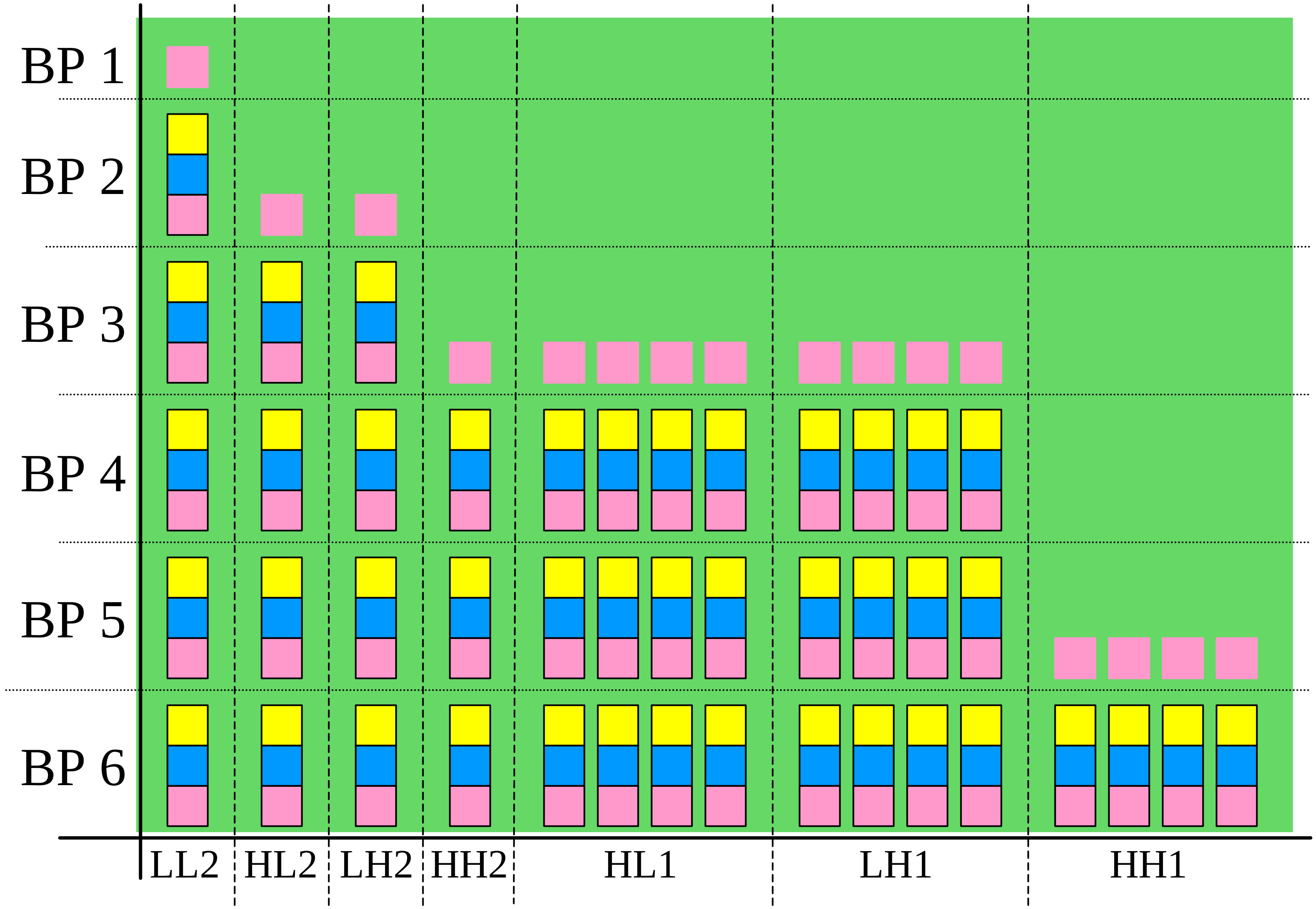
Lowest Resolution, Highest Quality



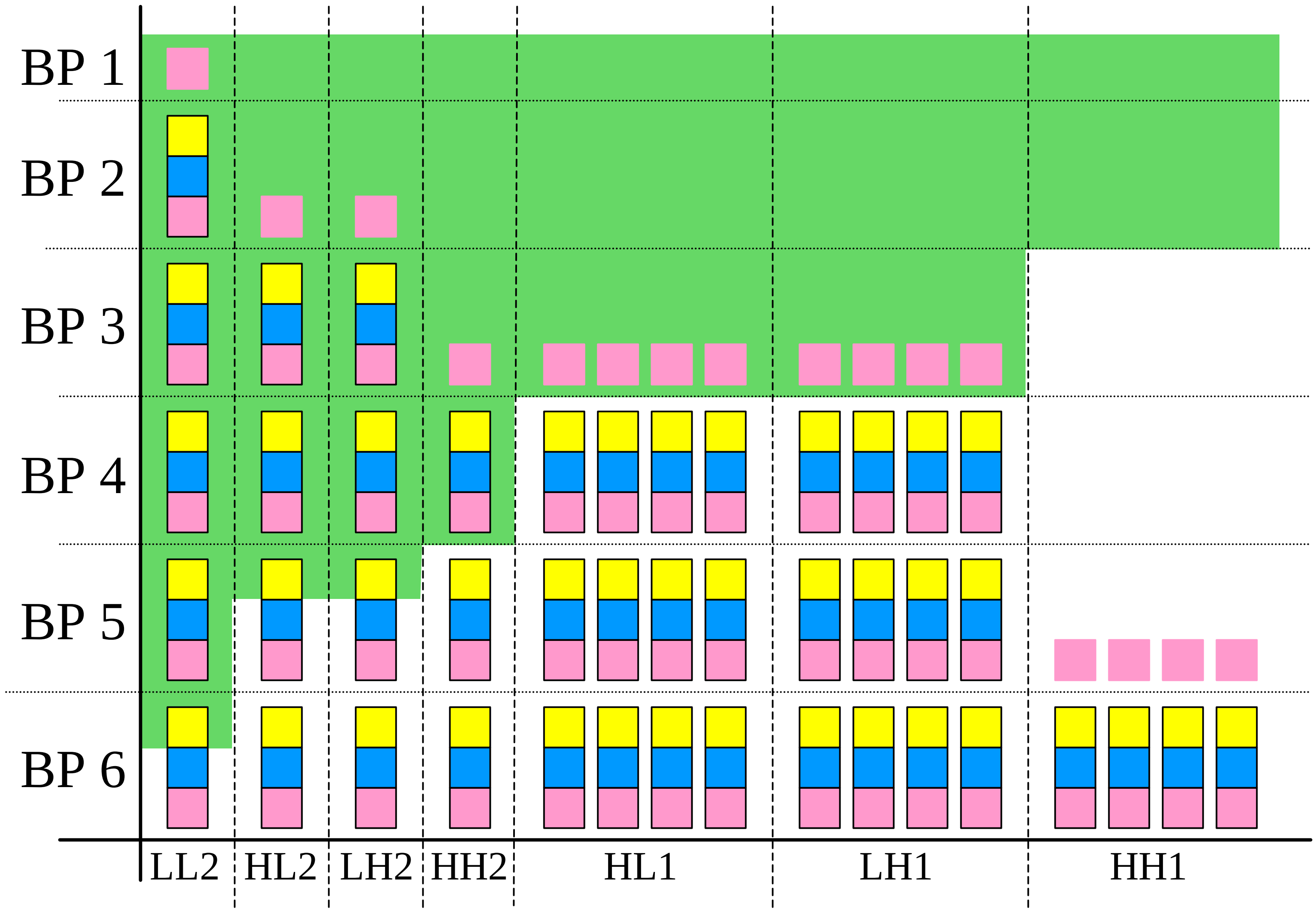
Medium Resolution, Highest Quality



Highest Resolution, Highest Quality

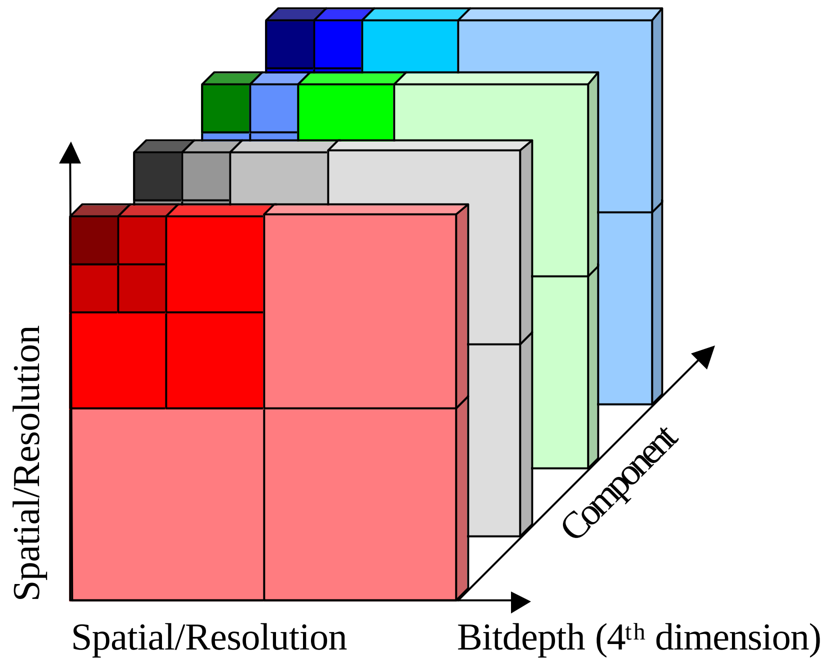


Highest Resolution, Target Visual Quality



JPEG 2000 Progression Types

- After wavelet processing, we have a four dimensional cube of data
  - Spatial/Resolution (two)
  - Component
  - Bitdepth
- JPEG 2000 allows progression along four dimensions
  - Layer (L)
  - Resolution (R)
  - Component (C)
  - Precinct or position (P)
- These are roughly equivalent as follows
  - Resolution & Precinct ⇔ Spatial/Resolution
  - Component ⇔ Component
  - Layer ⇔ Bitdepth



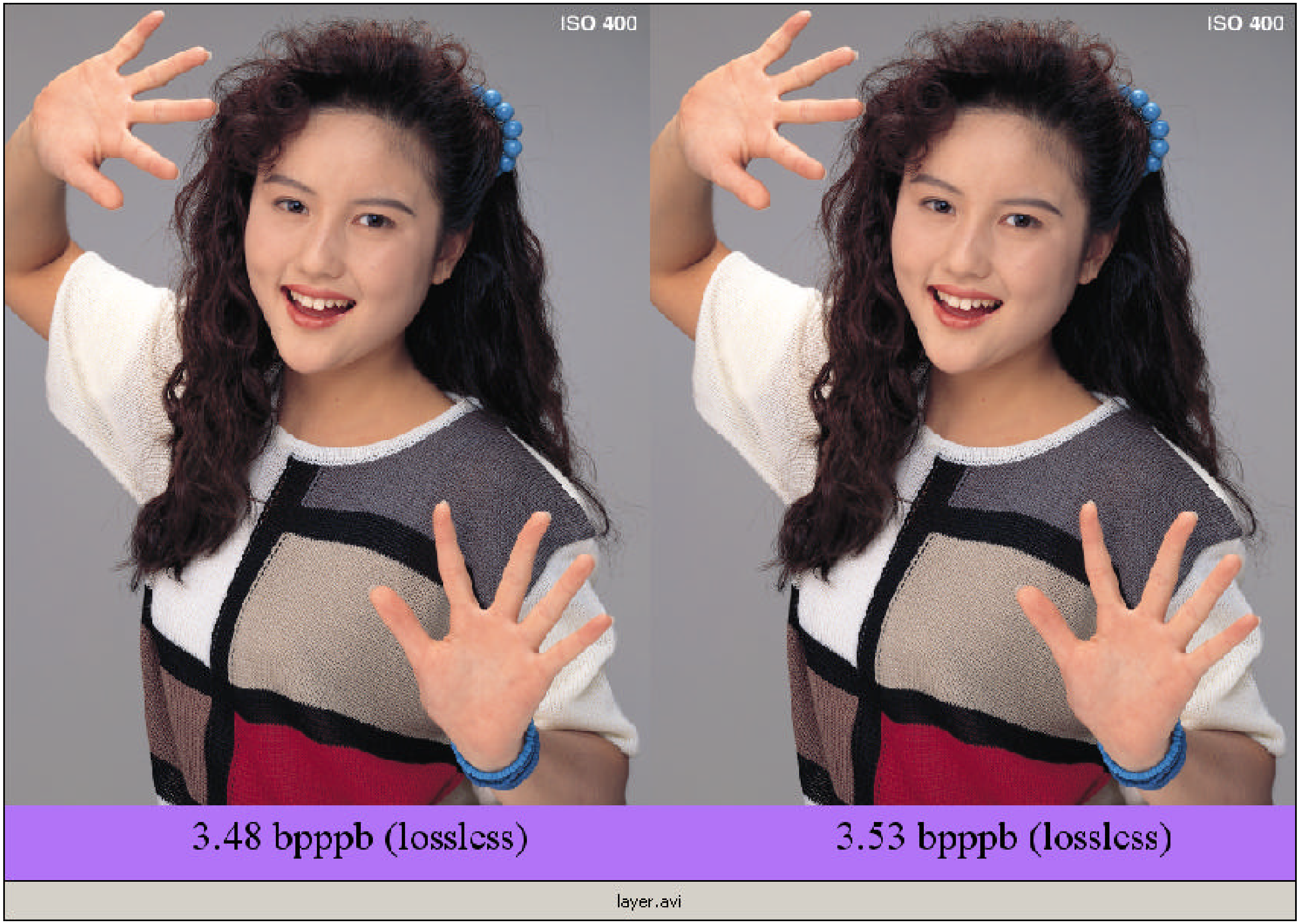
Wavelet processed components

JPEG 2000 Progression Types

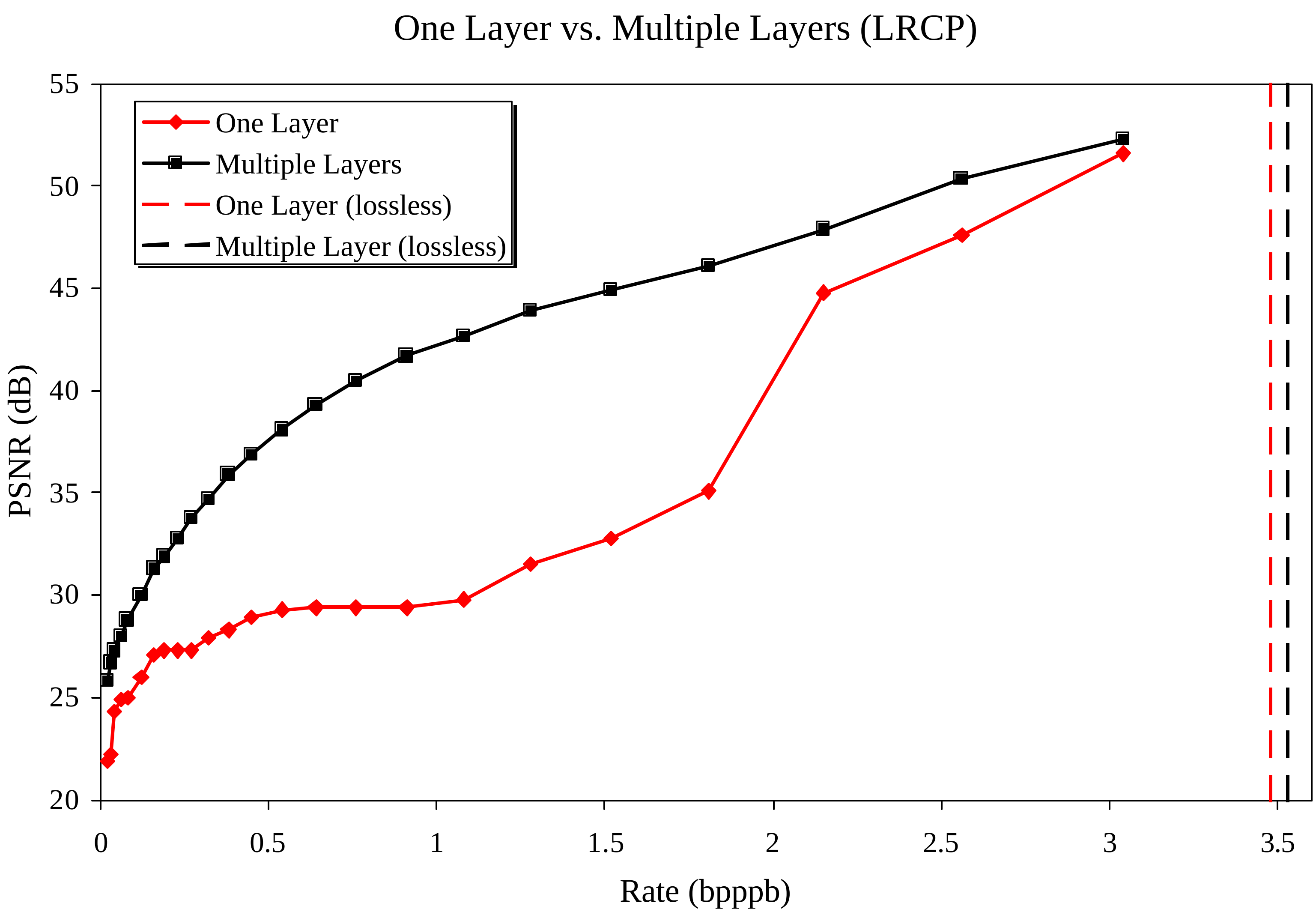
- There are five allowed progression types
  - LRCP, RLCP, RPCL, PCRL, CPRL
  - View progression as four nested loops, read left to right
- LRCP
  - Progression by SNR. Best full size image (SNR)
  - Loop over all layers
    - Loop over all resolution levels
      - Loop over all components
        - Loop over all precinct (positions)
- RLCP
  - Progression by resolution. Improving image quality for a fixed resolution
- Note that "C" (component) is in an inner loop in all progressions except CPRL

Effects of Layering

Single Layer | Multiple Layers



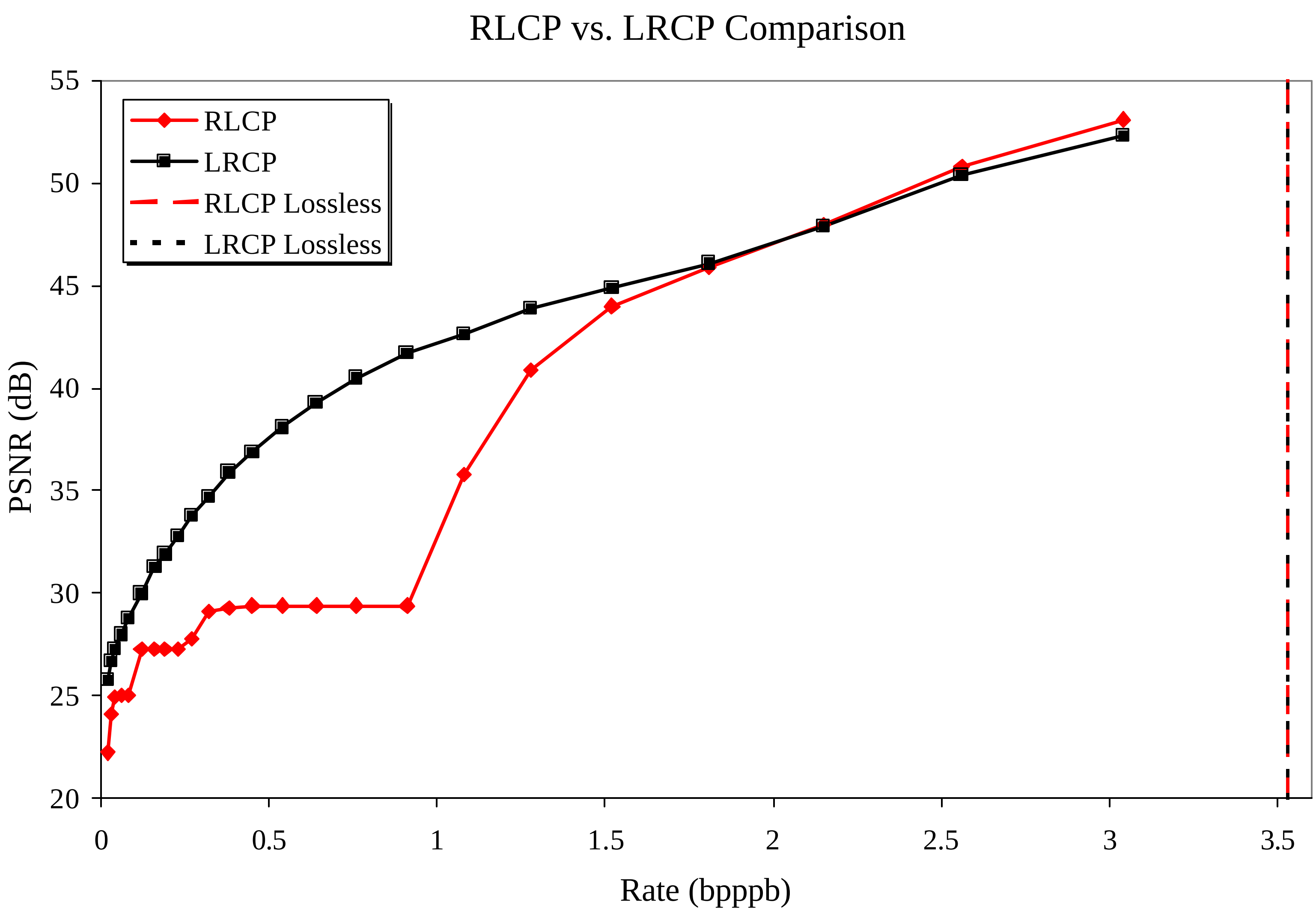
Effects of Layering



Effects of Progression

RLCP Progression | LRCP Progression

Effects of Progression

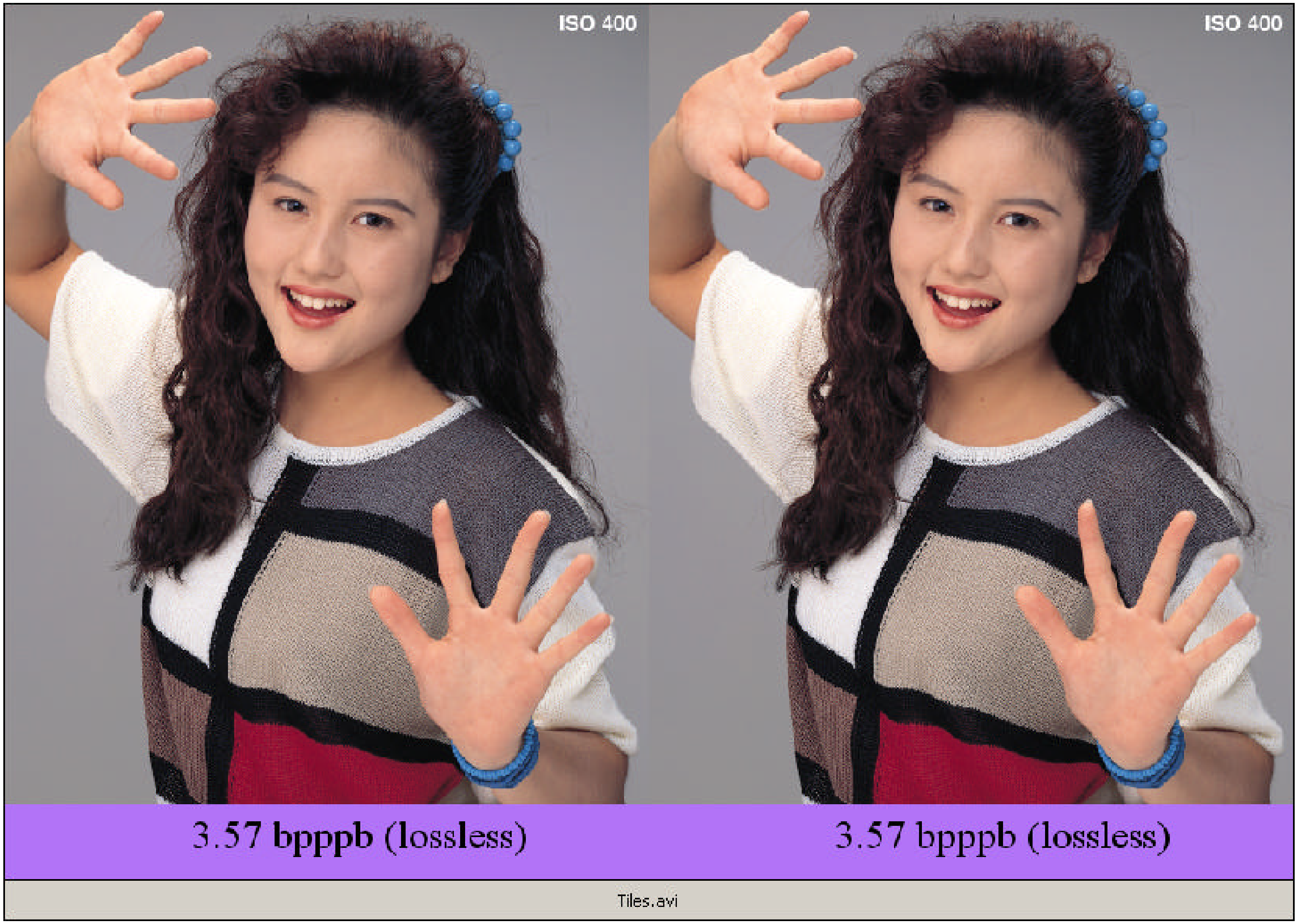


Tiles

- Tiles are independently coded sub images. Nothing crosses tile boundaries
  - Wavelet
  - Entropy coding
  - Layers
  - Progressions
- Tiles may be broken into tile parts. Tile parts from different tiles can be interspersed in a codestream
  - Only mechanism available to achieve “tile progression”
- In general, need to parse data out of tiles to achieve a different image quality
  - If all tiles are compressed at 2.0 bpp and you want 1.0 bpp, then need to go into each tile and get the 1.0 bpp

Tiled Image Parsing

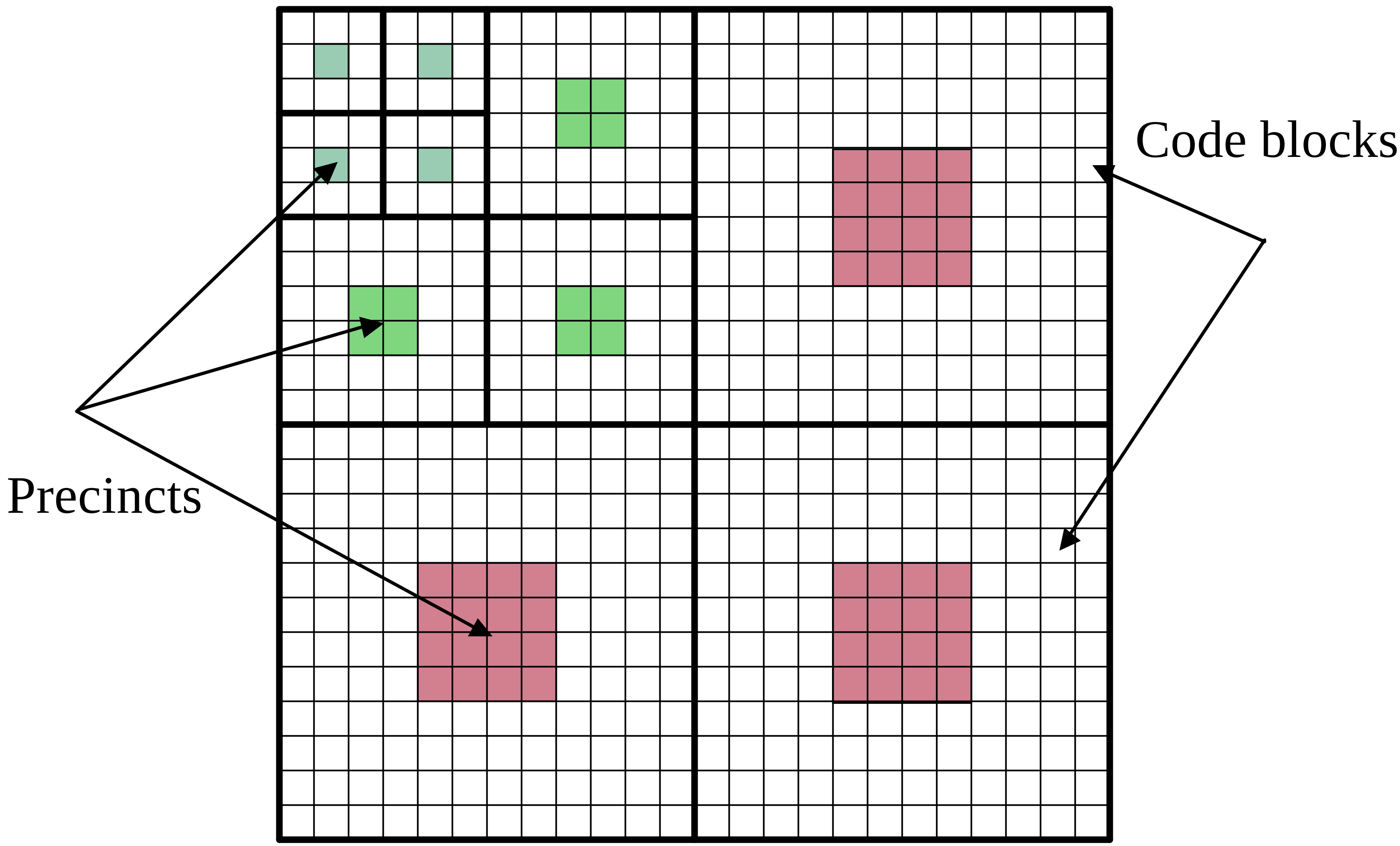
No Parsing Performed | Parsing Performed



JPEG2000 Bitstream Syntax

- **Precinct**: Each subband is divided into non-overlapping rectangles of equal size (except maybe at the boundaries where the size can be different) called precincts. Precincts provide some level of spatial locality in the bit stream and their boundaries are aligned with code blocks. Their size can vary for each tile, (color) component, and resolution.
- **Packet**: Consists of the compressed bit streams associated with a certain number of sub-bit planes from each codeblock in a precinct. Packets serve as one quality increment for one resolution level at one spatial location.
- **Layer**: is a collection of packets, one from each precinct at each resolution. It can be interpreted as one quality increment for the entire image at full resolution.

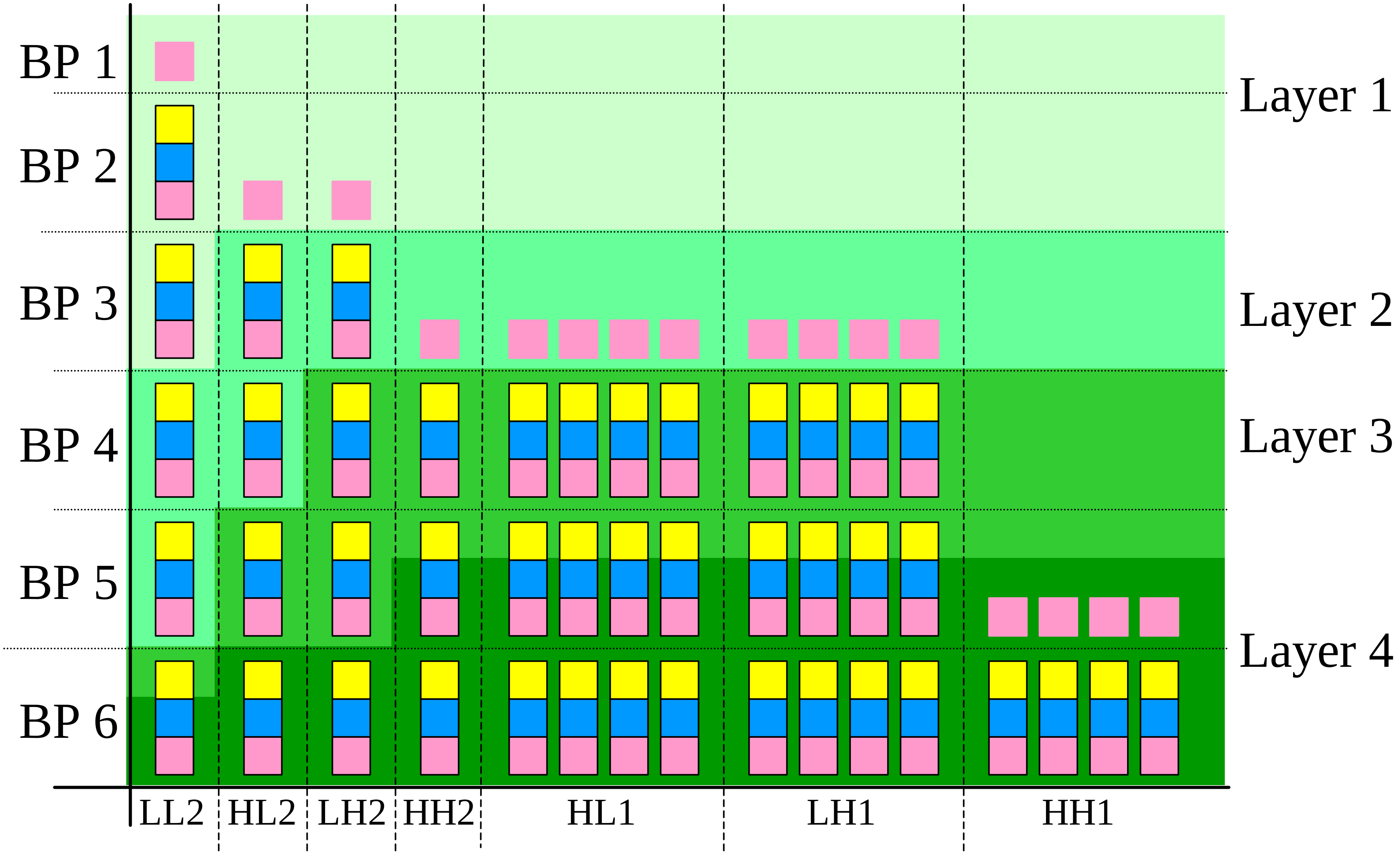
Example: Precincts and Codeblocks



Layers

- **Layer**: a collection of some consecutive bit-plane coding passes from all code-blocks in all subbands and components. Each code-block can contribute an arbitrary number of bit-plane coding passes to a layer.
- Each layer successively increases the image quality. Most often associated with SNR or visual quality levels.
- Layers are explicitly signalled and can be arbitrarily determined by the encoder
- The number of layers can range from 1 to 65535. Typically around 20. Larger numbers are intended for interactive sessions were each layer is generated depending on user feedback.

Bitstream Ordering with Four Quality Layers



Codestream

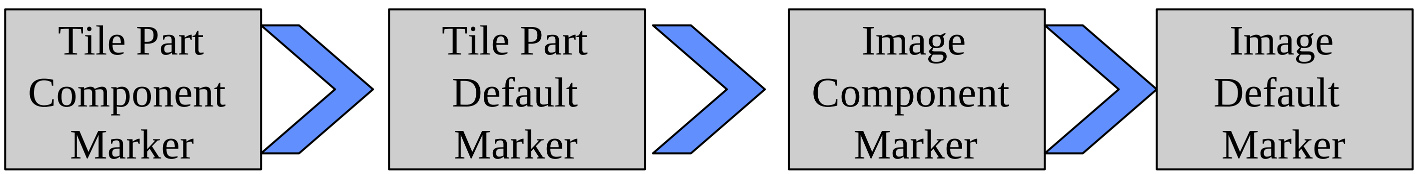
- **Codestream**: compressed image data with all the signaling required to properly decompress it.
- Composed of a main and tile headers, that specify coding parameters in a hierarchical way, plus the encoded data for each tile.
- The compressed data for a tile can be broken up in tile-parts, and the different tile-parts interleaved in the codestream to allow for non-tile progressiveness.
- The codestream is the minimum exchange format for JPEG 2000 encoded data, but usually the codestream is embedded in a file format.

JPEG 2000 Codestream Structure

- Delimiting markers
  - These are markers that are used in the start of most major sections of codestream and the very end.
  - Start of codestream (SOC), start of tile part (SOT), start of data (SOD), and end of codestream (EOC)
- Fixed information marker segments
  - This marker includes information that is required to properly decode the image.
  - Size Marker (SIZ) which includes reference grid size, tile size, resolution/sampling (relative to grid), image and tile offsets (into the grid), number of components, and component precision (data type and bit depth)

JPEG 2000 Codestream Structure

- Functional Marker Segments
  - These marker define the parameters used in the compression of a tile or an image. The order or precedence of the markers are



  - Default markers (can be used in image header or tile header)
    - coding style default (COD), quantization default (QCD), region of interest (RGN), progression order changes (POC)
  - Component Markers
    - coding style component(COC), quantization component (QCC)

JPEG 2000 Codestream Structure

- Functional Markers
  - Coding style default and coding style component include information on coding style, number of decompositions, code-block size, code-block style, wavelet transformation, precinct style
  - Quantization default and quantization component include information on the quantization for the derived, expounded or no quantization
  - Region of interest marker includes the location of the ROI
  - Progression order changes describes the bounds and progression order for any progression other than that in the COD marker in the main image header

JPEG 2000 Codestream Structure

- Pointer Marker Segments
  - Pointer markers are used for quick access to data that is required for decompression of a given location, resolution quality, or component.
  - All of the pointer segments define lengths of segments which allow fast rearranging of data and pointer markers
  - Main Header Markers
    - Tile part lengths (TLM), packet lengths main header (PLM), packed packet headers (PPM) (main header).
  - Tile part header
    - Packet length tile-part header (PLT), packed packet headers (tile part header) (PPT).

JPEG 2000 Codestream Structure

- In Bit stream markers
  - Start of packet (SOP) and end of packet (EOP) markers are used to isolate a given packet in a noisy environment
- Information marker segments
  - Component registration (CRG) marker is used to register components if the components do not have the same sampling to the reference grid
  - The comment marker (COM) is an open style marker that allows for unstructured data

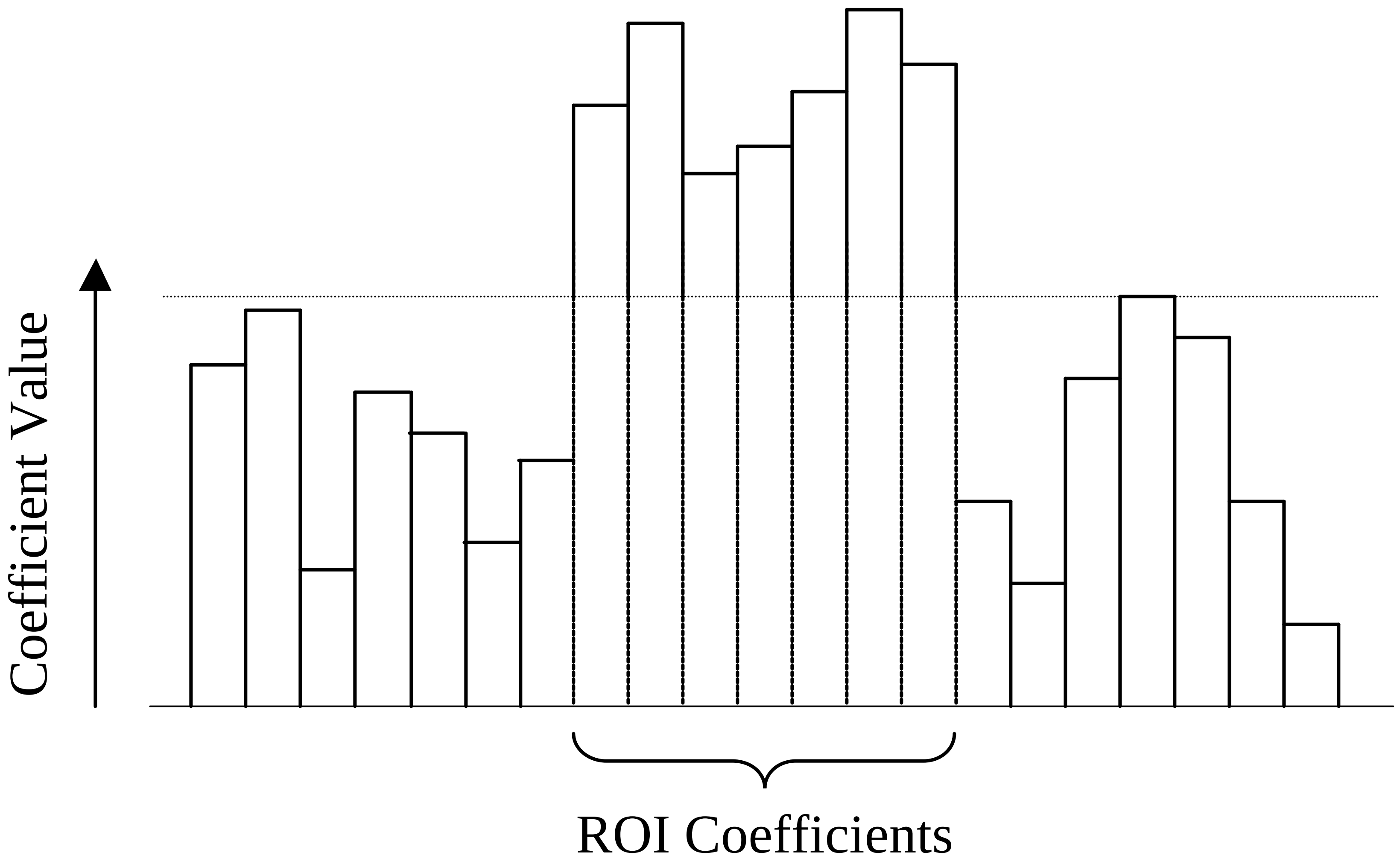
JPEG 2000 Codestream Structure

- The J2K codestream starts with the main header, followed by tile-part header(s), and bitstream(s) and ends with an EOC marker
  - The main header starts with SOC and SIZ markers, then followed (in any order) by COD and QCD markers and possibly QCC, RGN, POC, PPM, TLM, PLM, CRG, COM
  - The tile part header starts with SOT marker and finishes with SOD and can contain (in any order) COD, COC, QCD, QCC, RGN, POC, PPT, PLT, COM
  - The bitstream may contain SOP and EPH markers

ROI Coding in Part I: Maxshift Method

- In the **Maxshift** method, the wavelet coefficients in the ROI region are scaled up by a fixed number of bits *s*, so that the smallest shifted nonzero ROI coefficient is larger than the largest BG coefficient. The parameter *s* is signaled in the bit stream. As a result, the decoder can discriminate between the ROI and BG coefficients by comparing each decoded value to a threshold. Pros and cons are:
  - It allows for multiple regions of arbitrary shape ROI without the inclusion of the shape information in the bit stream or need for ROI mask generation at the decoder.
  - The user can prioritize the coding of the ROI region over the BG but does not have control over the quality differential between ROI and BG.

Maxshift Method of ROI Coding



Current JPEG 2000 Commercial Support

Commercial Software, IP, and Silicon

Kodak

JPEG 2000 Software

AWARE

Power image

algoVision LuraTech

University of Victoria

DSPG@UVic
Digital Signal Processing Group

Kakadu

JPEG 2000 Firmware

ALMA TECHNOLOGIES

SYNOPSYS

Adelante Technologies

AWARE

AMPHION

Power image

ateme
l'oasis des technologies

JPEG 2000 Silicon

ANALOG DEVICES

WIS TECHNOLOGIES

RICOH
Image Communication

Commercial Products Supporting JPEG 2000

Kodak

QuickTime

Nikon instruments

Leica Geosystems

YAHOO! Messenger

JAVA

intel.

SONY

Kodak

Adobe

NEC

Snowbound software

CORELDRAW GRAPHICS SUITE

NCTsoft

FRONTIER MAPPING

PEGASUS IMAGING CORPORATION

SANYO

GlobeXplorer

PICTURE ELEMENTS INCORPORATED

ALPARY SOFT

Pixel Translations ISIS THE STANDARD FOR IMAGING TOOLS

CSmap Desktop Mapping

SE SEAKR Engineering Incorporated

MicroImages microimages.com

MATROX IMAGING

MAPQUEST

eyespy EyeSpy Pictures That Fly.

LEAD TECHNOLOGIES

TEXAS INSTRUMENTS

AGI

ANYTHING3D.com

CAST

PHOTODEX CORPORATION

Netscape 7.0

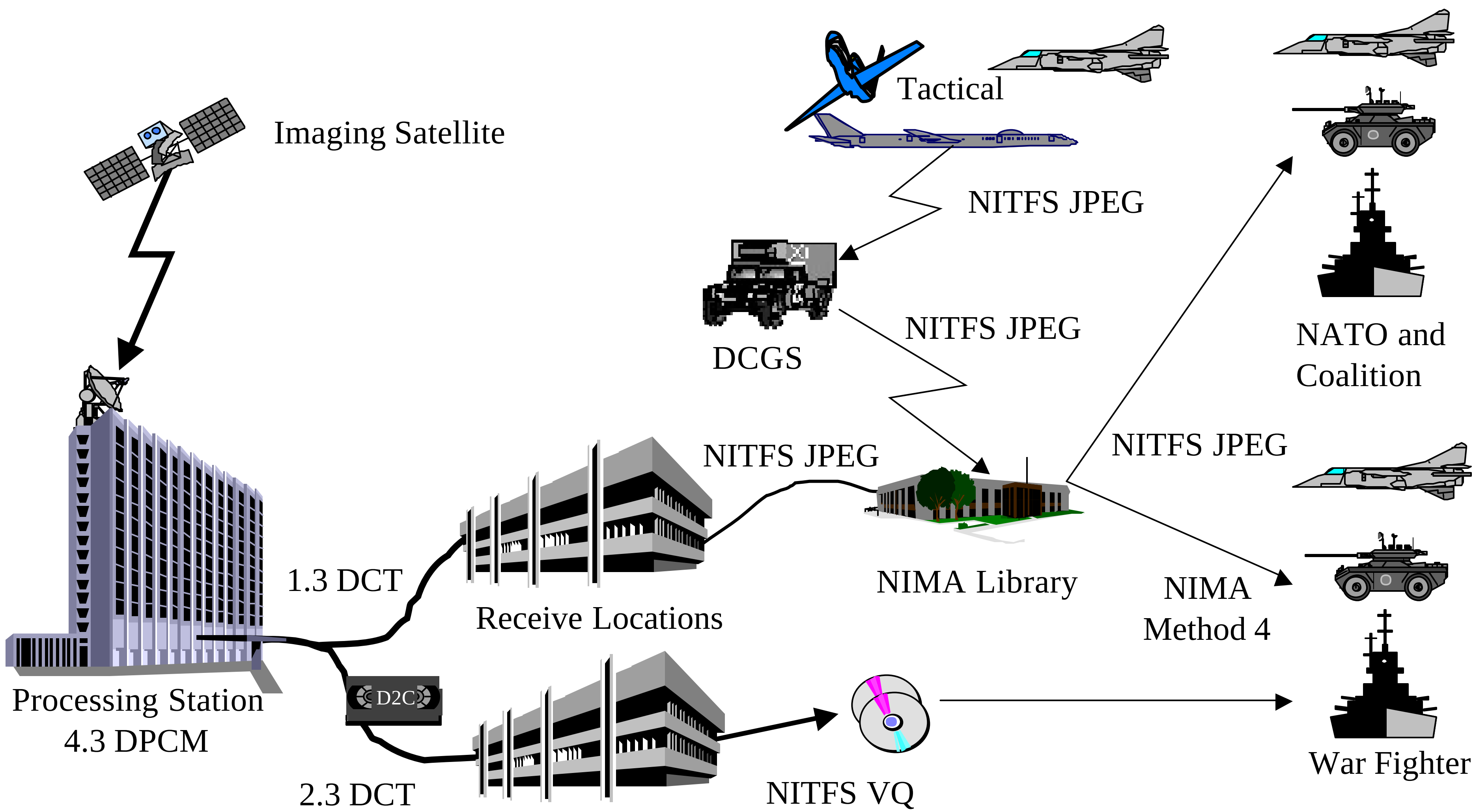
Microsoft Internet Explorer

Windows Media 9 Series

Current Dissemination of Imagery

What and how compression is used in today's NSGI system.

United States Imagery and Geospatial Information Service

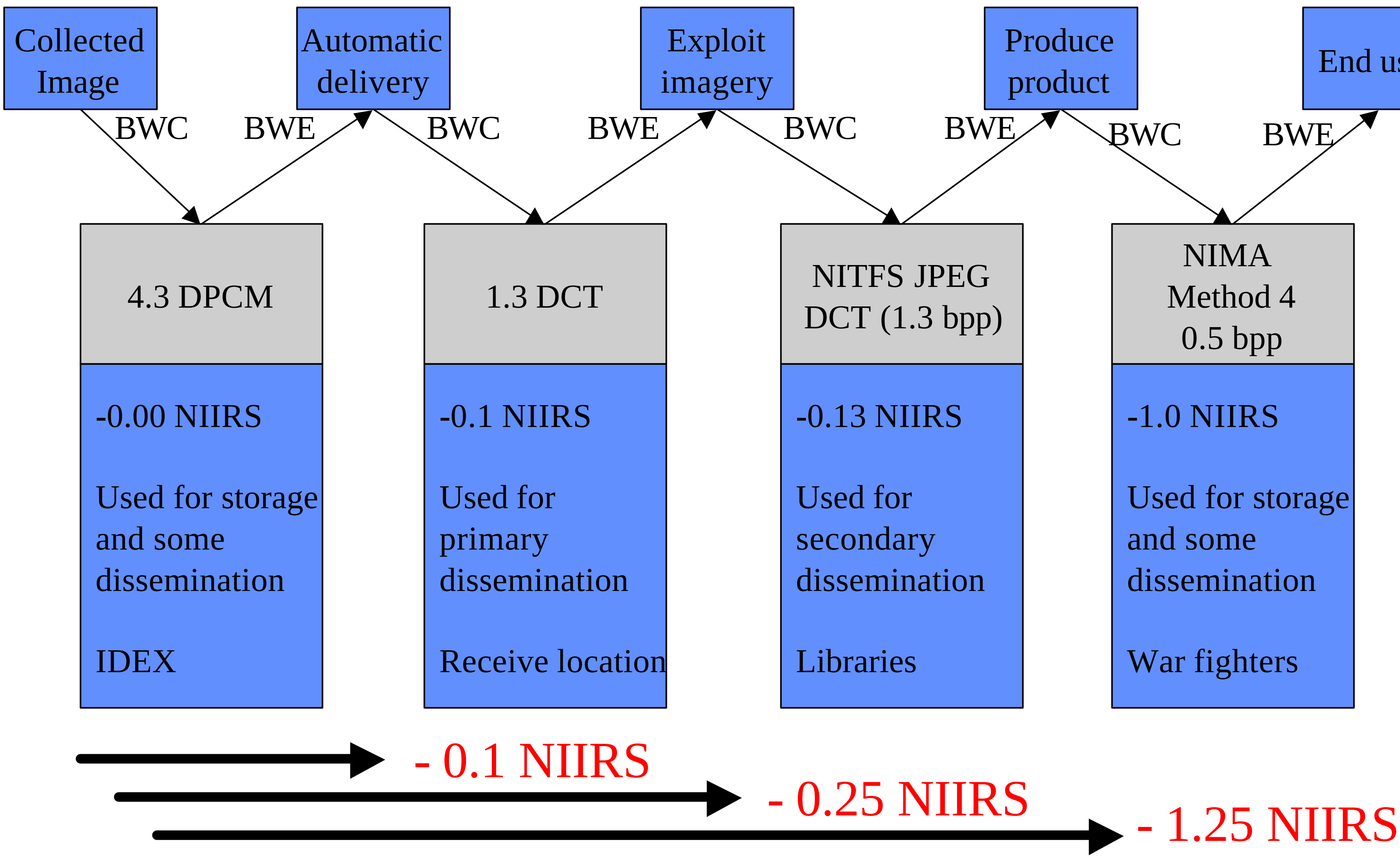


Current USIGS Compression of Imagery

- The primary system's compression algorithms were developed for high quality for primary dissemination and exploitation of data.
- The secondary system's compression algorithm was adopted from commercial sources because of flexibility and COTS availability.
- The tactical BWC was derived from JPEG to be backwards compatible and meet the tactical dissemination requirements.

4.3 DPCM	1.3 DCT	NITFS JPEG DCT (1.3 bpp)	NIMA Method 4 0.5 bpp
-0.00 NIIRS	-0.1 NIIRS	-0.13 NIIRS	-1.0 NIIRS
Used for storage and some dissemination	Used for primary dissemination	Used for secondary dissemination	Used for storage and some dissemination
IDEX	Receive location	Libraries	War fighters
Primary Users		Secondary Users	Tactical Users

USIGS Dissemination of Imagery



J2K Profile General Recommendations

- J2K Tiger team defined JPEG 2000 recommended best practices
  - Profile defines parameters for best quality and most functionality for the USIGS architecture
- Recommendations for best image quality results
  - Select the 5-3 numerically lossless for radiometric images (IR,MS, HSI)
  - Defined the usage of the 9-7 visually lossless (Pan, SAR)
- To enable quality scalability
  - 19 truncation quality truncation points to support all applications
  - Lossless (radiometric exploitation), 3.5– 2.0 bpp (MC&G), 2.0 – 1.0 bpp (1st phase visual exploitation), 1.5 – 0.5bpp (Tactical users), and 0.5 – 0.03125 (bandwidth constrained users)
- To enable resolution scalability
  - 5 Levels of wavelet decompositions (R0 – R6 available)
- To enable fast access chipping
  - Tiles of 1024-by-1024 with tile offset in the file header
  - TLM and PLT markers for pointing to chips and quality layers

Compression Overview

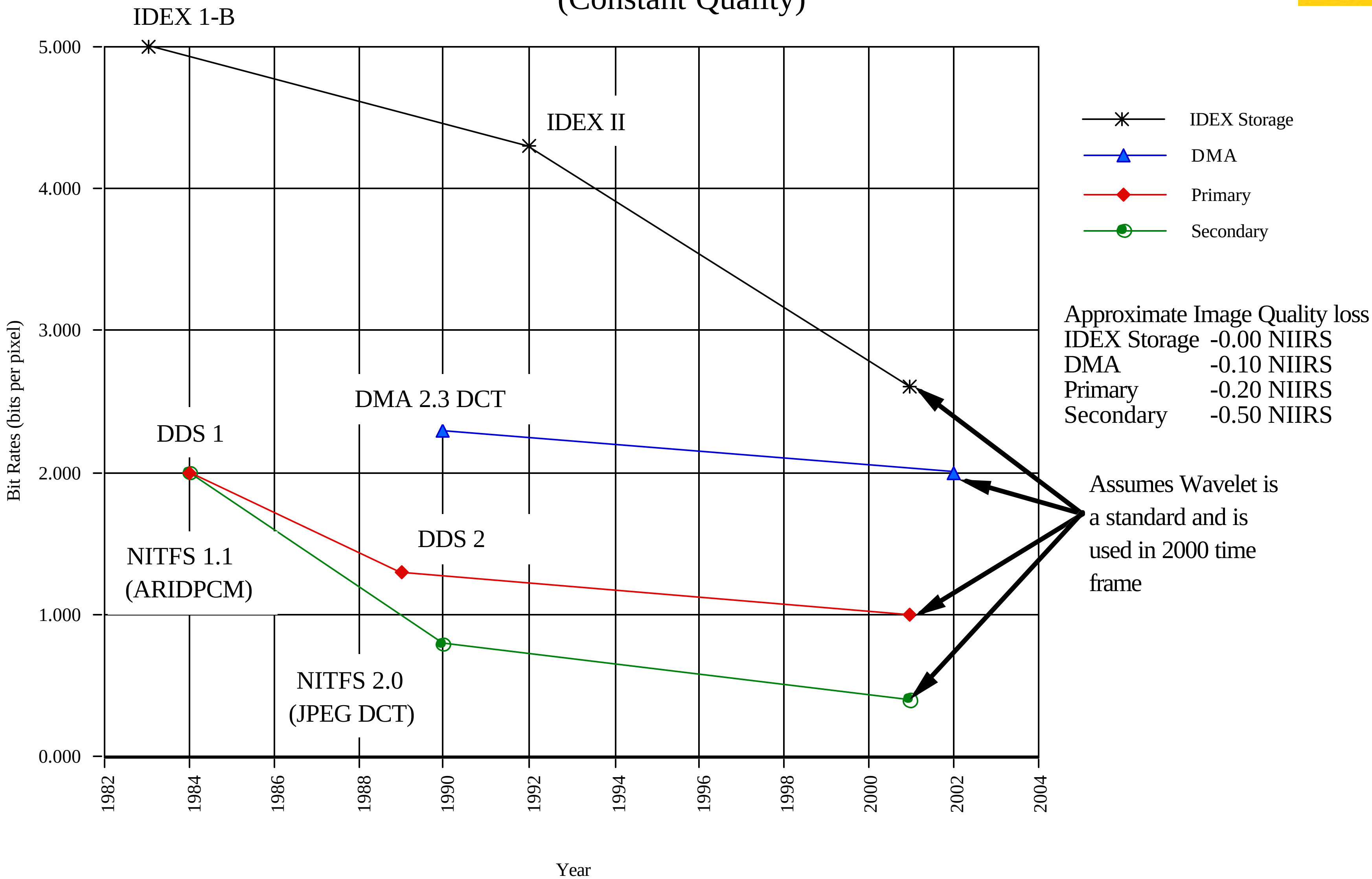
Compression Algorithm	Transform Technique	Quantization	Encoding	Comments
4.3 DPCM	Linear prediction from neighboring pixels	Table look-up	Variable length Huffman encoding	• Low complexity • High quality • Low compression ratio
NITFS JPEG DCT	8-by-8 block Discrete Cosine Transform (DCT)	Human Visual System (HVS) Response quantization in DCT space	Variable length Huffman encoding	• Can be rate controlled • 8-by-8 transform used for speed
1.3 / 2.3 DCT	32-by-32 block DCT	HVS quantization	Variable length Huffman encoding	• Rate controlled to either 1.3 bpp or 2.3 bpp
Vector Quantization	No transform performed	Vector code book matching	Code book numbers	• Low channel error susceptibility
NIMA Method 4	Down-sampled followed by JPEG DCT	JPEG DCT	JPEG DCT	• Achieves very low bit rate at reasonable quality
JPEG 2000	Wavelet-based sub-band transform	Scalar Quantization with Dead-Zone	Bit-Plane Arithmetic encoder	• Highest quality of any of the algorithms • Most functional

Compression Overview

Algorithm	Advantage	Disadvantage
4.3 DPCM	• Low complexity (low power/size/weight) • Visually lossless quality • Low memory requirements • Government standard • Rate controlled	• Low compression ratios compared to frequency-based transform techniques.
1.3 DCT	• High quality • Military standard • Rate controlled	• High complexity (32-by-32 transform, rate-control) • Blocking artifacts
2.3 DCT	• Near lossless quality • Government standard • Rate controlled	• High complexity (32-by-32 transform, rate-control) • High bit rate (2.3 bpp)
NITFS JPEG DCT	• International/commercial standard • Low cost implementation (COTS) • Low complexity (8-by-8 transform)	• Blocking artifacts • Lower quality than 1.3 DCT and wavelets
NITFS VQ	• Low complexity for decompression • Low susceptibility to channel error • High quality on DMA maps • Military standard	• High complexity for compression • Relatively poor quality on images
NIMA Method 4	• Interoperable with NITFS JPEG • High quality at low bit rates • Military standard	• Does not perform well at higher bit rates • Not flexible
Wavelets	• Better quality to compression ratio than any other compression algorithm • Significantly more functionality • Commercial Standard	• Large memory requirements • Computational Complexity • Significant start-up cost

Compression Rate Improvement Over Time

(Constant Quality)



Related Issues (File Formats)

- TFRD
  - Used for primary dissemination
  - End users exploit imagery to develop image products
    - Requirements included SDE information
  - Does not include graphics or overlays
- NITFS
  - Used for secondary dissemination of exploited data
  - End users only use imagery and do not exploit imagery
  - Require graphics and overlays
- The two are trying to extend their file format and compression to incorporate the others needs.
  - NIMA now has control over all of the systems (primary/secondary)

Acronym List

- Image Processing
  - DRA = Dynamic Range Adjustment
  - TTC = Tone Transfer Curve
  - RRDS = Reduced Resolution Data Sets
  - MSE = Mean Square Error
  - NIIRS = National Image Interpretability Rating Scale
  - RMSE = Root MSE
  - Color space transformations
    - Red Green Blue (RGB) Hue Intensity Transform (YCrCb, YIQ, YUV)
  - HVS = Human Visual System
- General
  - COTS = Commercial off the self
  - GSD = Ground Sampled Distance
  - DPI = Dots Per Inch
  - IA = Image Analyst

Acronym List

- Compression
  - DPCM = Differential Pulse Code Modulation
  - DCT = Discrete Cosine Transform
  - JPEG = Joint Photographic Experts Group
  - MPEG = Motion Picture Experts Group
  - VQ = Vector Quantization
- Systems/ users/ formats
  - USIGS = United States Image and Geospatial System
  - DDS = Defense Dissemination System
  - NITFS = National Imagery Transmission Format Standards
  - NIMA = Nation Imagery and Mapping Agency
  - TFRD = Tape Format Requirements Document
  - RE/RL = Receive Entity
  - DE = Distribution Entity
  - DoD = Department of Defense