Wavelet-based Image Compression

Situations where image compression offers a solution

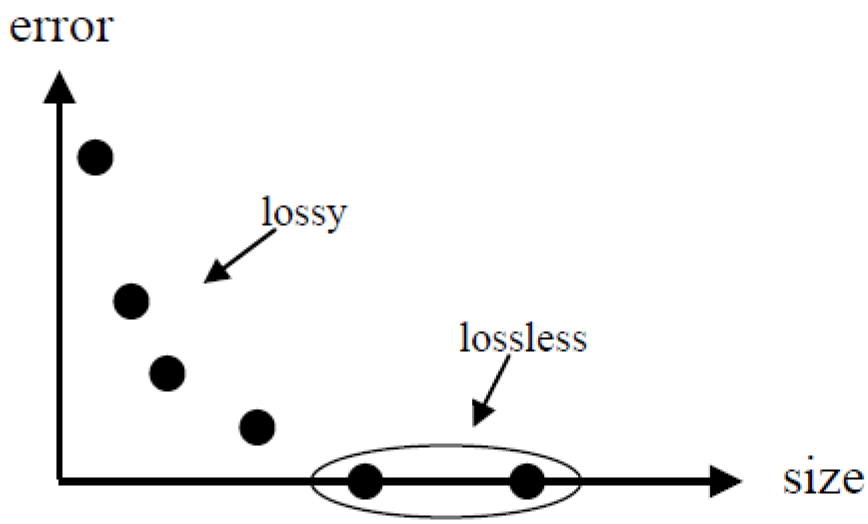
- Video
  - 480p with 10 key frames/sec requires 0.51 GB for one minute.
- Digital cameras
  - 1 MP and 8 MP images require 3 MB and 22.8 MB of storage.
- Reducing storage and transmission costs lead to image compression.

Overview

- Redundancies in images.
- What typical image compression systems do.
- Current methods for natural image compression.
- State-of-the-art image compression with wavelet representations.

Representing information with data

- More data is usually used than is absolutely required.
- Reduce data by minimizing redundancy and/or reducing detail.



Coding redundancy

- Some values are more common than others.
- Example, this image has four colors:

Huffman tree construction

Original source		Source reduction			
Symbol	Probability	1	2	3	4
a_2	0.4	0.4	0.4	0.4	0.6
a_6	0.3	0.3	0.3	0.3	0.4
a_1	0.1	0.1	0.2	0.3	
a_4	0.1	0.1	0.1		
a_3	0.06	0.1			
a_5	0.04				

Image from “The Data Compression Book” by Mark Nelson

Huffman coding

- Average bits per symbol with BCD: 3
- New average bits per symbol: 2.7
- Redundancy has been decreased by 10%.

r_k	$p_r(r_k)$	*Code 1*	$l_1(r_k)$	*Code 2*	$l_2(r_k)$
$r_0 = 0$	0.19	000	3	11	2
$r_1 = 1/7$	0.25	001	3	01	2
$r_2 = 2/7$	0.21	010	3	10	2
$r_3 = 3/7$	0.16	011	3	001	3
$r_4 = 4/7$	0.08	100	3	0001	4
$r_5 = 5/7$	0.06	101	3	00001	5
$r_6 = 6/7$	0.03	110	3	000001	6
$r_7 = 1$	0.02	111	3	000000	6

Image from “The Data Compression Book” by Mark Nelson

Arithmetic coding

- $a_1 = 0.2$, $a_2 = 0.2$, $a_3 = 0.4$, $a_4 = 0.2$
- $a_1\ a_2\ a_3\ a_3\ a_4$ can be encoded as 0.068
- 3/5 decimal digits per symbol.

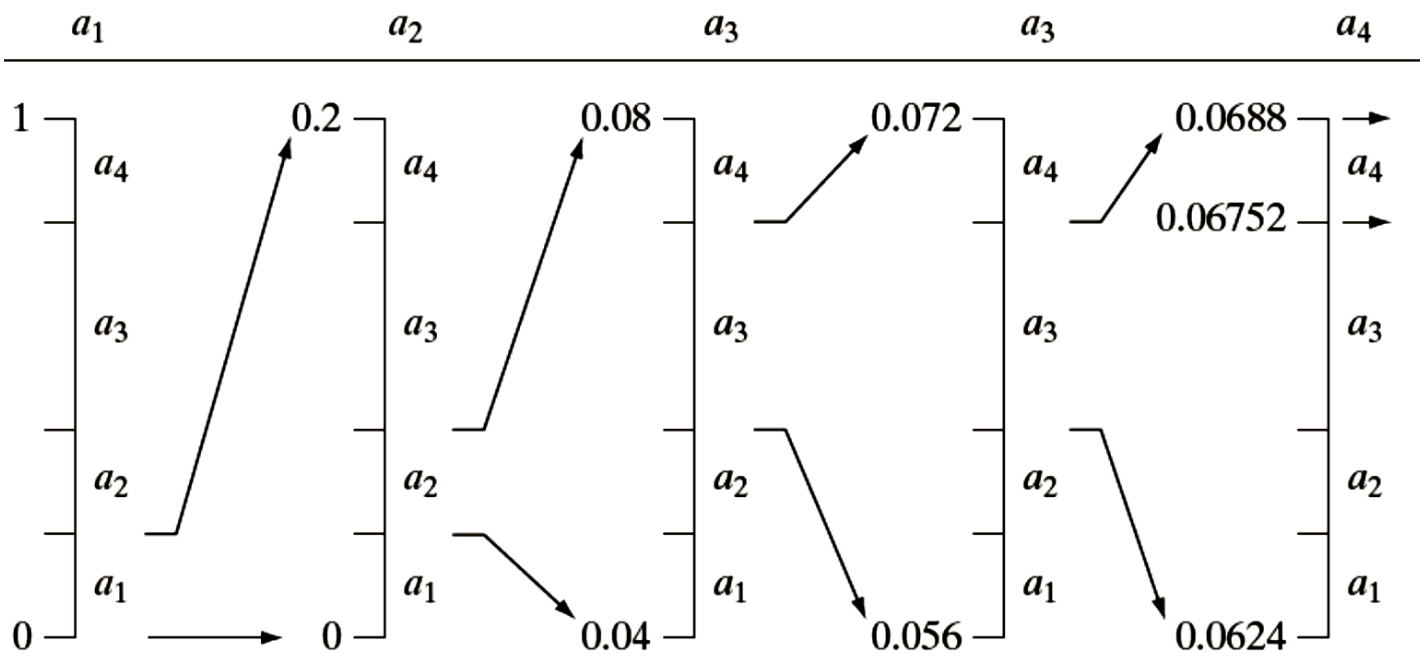


Image from “The Data Compression Book” by Mark Nelson

Inter-pixel redundancy

- Adjacent pixels tend to be correlated.
- Transform to reduce correlation.
  - Transforms do not provide any compression.
  - Transforms are reversible mappings between domains.

Psycho-visual redundancy (lossy)

- We perceive substantial variances in intensity.
  - But miss minor ones.

Image from http://www.cse.unr.edu/~bebis/CS474/

Psycho-visual redundancy (lossy)

- The human eye is less sensitive to chroma.
  - Only applicable in natural images.
  - Decompose RGB into luminance and chroma.
  - Chroma is often down sampled (4:2:2 or 4:1:1).
  - Detail reduction is often imperceptible.
  - Formula for lossless YUV:

$$\begin{pmatrix} Y_r \\ U_r \\ V_r \end{pmatrix} = \begin{pmatrix} \left\lfloor \dfrac{R+2G+B}{4} \right\rfloor \\ R-G \\ B-G \end{pmatrix}$$

Original image

Image by Haneburger published on Geolocation.ws, filtered by demo program.

RGB channels

Image by Haneburger published on Geolocation.ws, filtered by demo program.

YUV decomposition

Image by Haneburger published on Geolocation.ws, filtered by demo program.

Criteria for compression

1. What is the amount of information in an image?
2. How much redundancy can be removed?
3. What is the minimum amount of data required for an "adequate" reconstruction?
   – Visual quality measured by subjective and objective fidelity.
4. How computationally complex is the encoder?
   – Modern techniques were previously unfeasible.

Anatomy of an image compressor

- Image compressors use the following steps to tackle the various redundancies:
  - Transformation
    - Inter-pixel redundancy
  - Chroma decomposition and quantization
    - Psycho-visual redundancy
  - Symbol coder
    - Coding redundancy
- The output from the symbol coder is then transmitted.

Discrete Cosine Transform

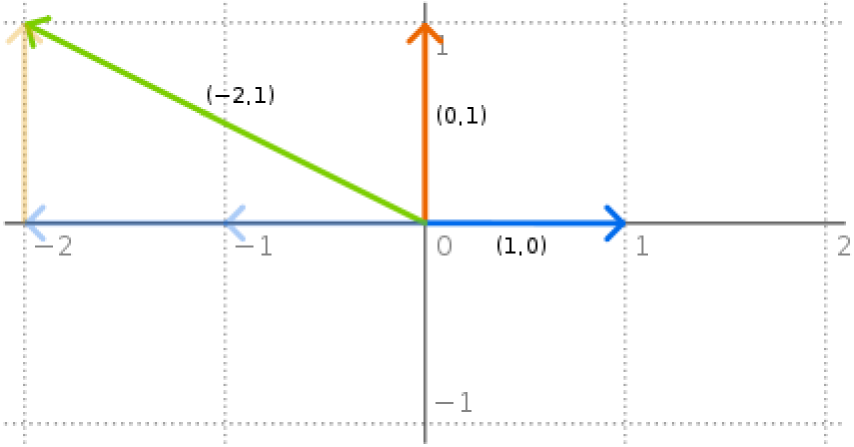
- Current standard transform for minimizing inter-pixel redundancy.
- Decomposes a block into a weighted sum of sinusoidal waves of different frequencies.
  - Transforms pixels from image space into frequency space.
- Discrete real-valued descendant of the Fourier Transform used in signal processing.

Basis functions

- Definition similar to Linear Algebra, but with functions instead of vectors.
- A basis is a set of functions with the following properties:
  - Any weighted sum of the basis functions can represent every possible outcome in a given space.
  - Every pair of functions within the basis is linearly independent.
- Basis for a N-dimensional space has N entries.

Analogy with basis vectors

- The orange and blue vectors form a basis for the surface.
- The green and yellow lines can be expressed as weighted sums of the basis vectors. Their representation is a pair of weights for the basis vectors.



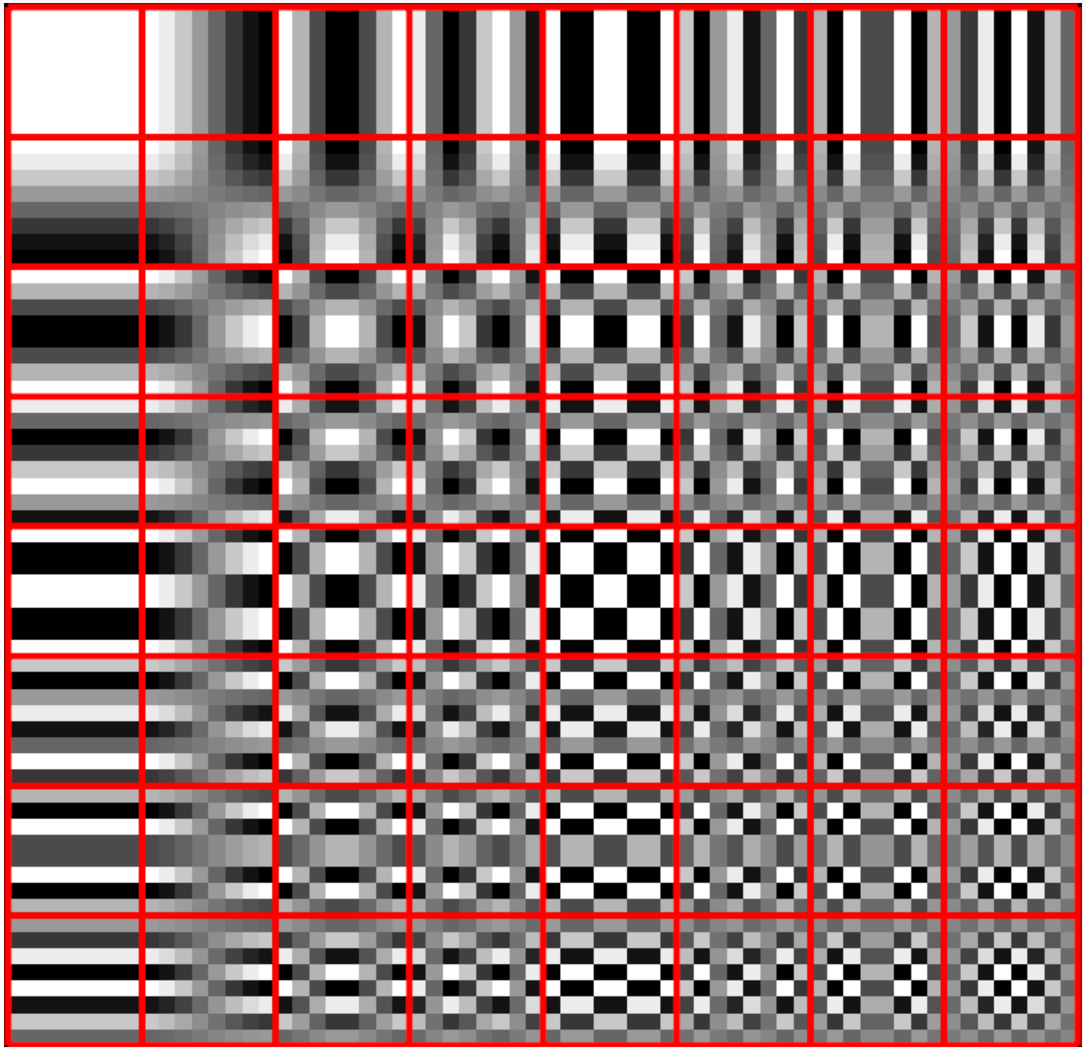
Basis functions

- The basis functions for a fixed-size DCT transform are calculated from a formula, not an image.
- The transform only finds the weights associated with the basis functions such that their weighted sum will reproduce the original image block.

DCT transform

- JPEG uses 8 frequencies in two dimensions that are used to represent each 8x8 block.
  - 64 pixels are used to calculate 64 coefficients which represent the energy present at each frequency.
  - This separates approximation and detail information.
  - The weights are stored as DCT coefficients.

DCT basis



DCT basis generated by demo program.

DCT applied to an image

DCT applied to an image

DCT calculated by demo program.

Quantization

- Real-valued coefficients in the frequency domain can be quantized to a discrete set of values.
- Coefficients for higher frequencies are less important.
- Each DCT coefficient is divided by the corresponding entry in the quantization matrix.
- Quantization matrices are determined subjectively.

Quantization

- Here is an example of quantizing a value of 21.82 with a step size of 10. Reconstruction would give -2 * 10 = -20.

$$\text{Quantizer Index} = -\left\lfloor \frac{21.82}{10} \right\rfloor = -2$$

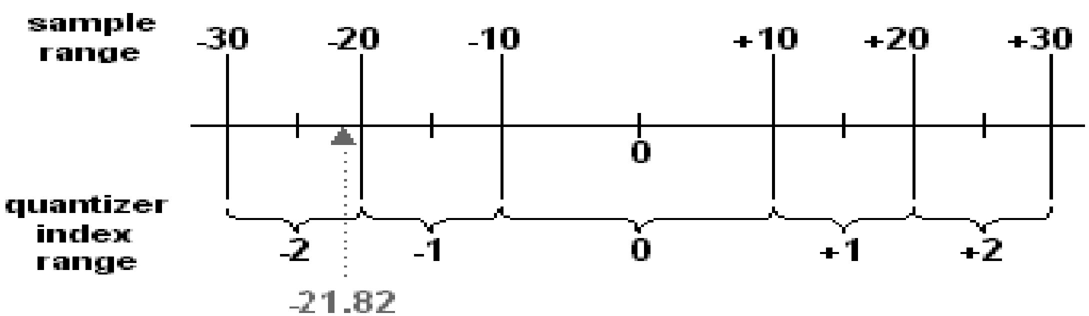


Image from the JPEG 2000 specification.

Quantization matrix

- Note that coefficients with greater significance do not lose too much detail.

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

Quantization matrix from Photoshop’s manual.

Storage

- The coefficients can be traced with a zig-zag walk and run-length encoded.
- All surviving coefficients are written out in order using statistical coding.

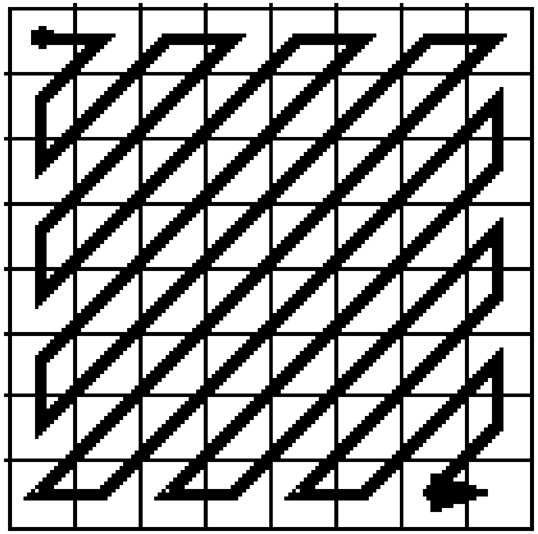


Image from the JPEG 2000 specification.

Disadvantages associated with the DCT

- No global detail reduction.
- Blocking artifacts caused by discontinuities.
  – Varying quantization errors at edges.
- Ringing artifacts caused by Gibb's Phenomenon.
  – Finite sum of sine waves overshoot edges.

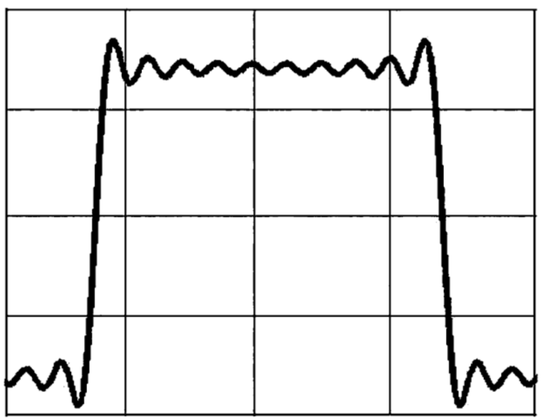


Image from http://people.clarkson.edu/~ajerri/books/

Disadvantages associated with the DCT

Image from http://en.wikipedia.org/wiki/File:Phalaenopsis_JPEG.png

Wavelets

- A waveform with limited duration and average value of zero.
- A wavelet's position and scale make it useful for joint frequency-time analysis.
- Wavelets are not restricted to sinusoidal waves.

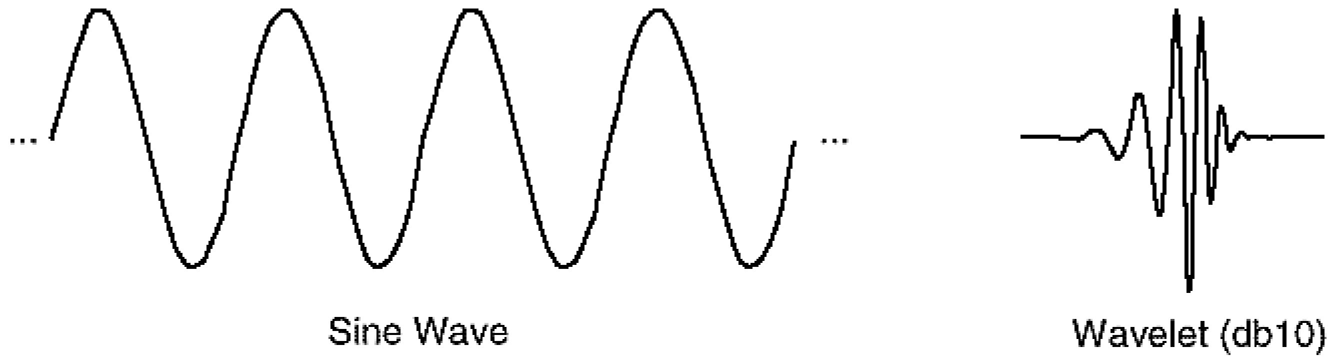


Image from http://klapetek.cz/wdwt.html

Global reduction with wavelets

- The image on the right was reconstructed with only 2% of its most significant coefficients.

Image from Matlab's DWT test cases.

Local versus global reduction

- The equivalent after DCT with only 2% of its top-left coefficients.

Image from Matlab's DWT test cases.

Discrete Wavelet Transform

- Digital images are discrete.
- Therefore, there is a finite number of coefficients for a set of shifted and scaled wavelets.
- Multi-resolution analysis with a dyadic decomposition is used to separate high-frequency details from low-frequency approximations.

Multi-resolution analysis

- Running the input through high-pass and low-pass filters produces detail and approximation information.
  - Each produces half the number of points in the input.
- The approximation information is now at half of the original resolution.
  - The analysis repeated at this “new” resolution.
  - Effectively scales wavelets to twice their original width during the next pass.

Multi-resolution analysis

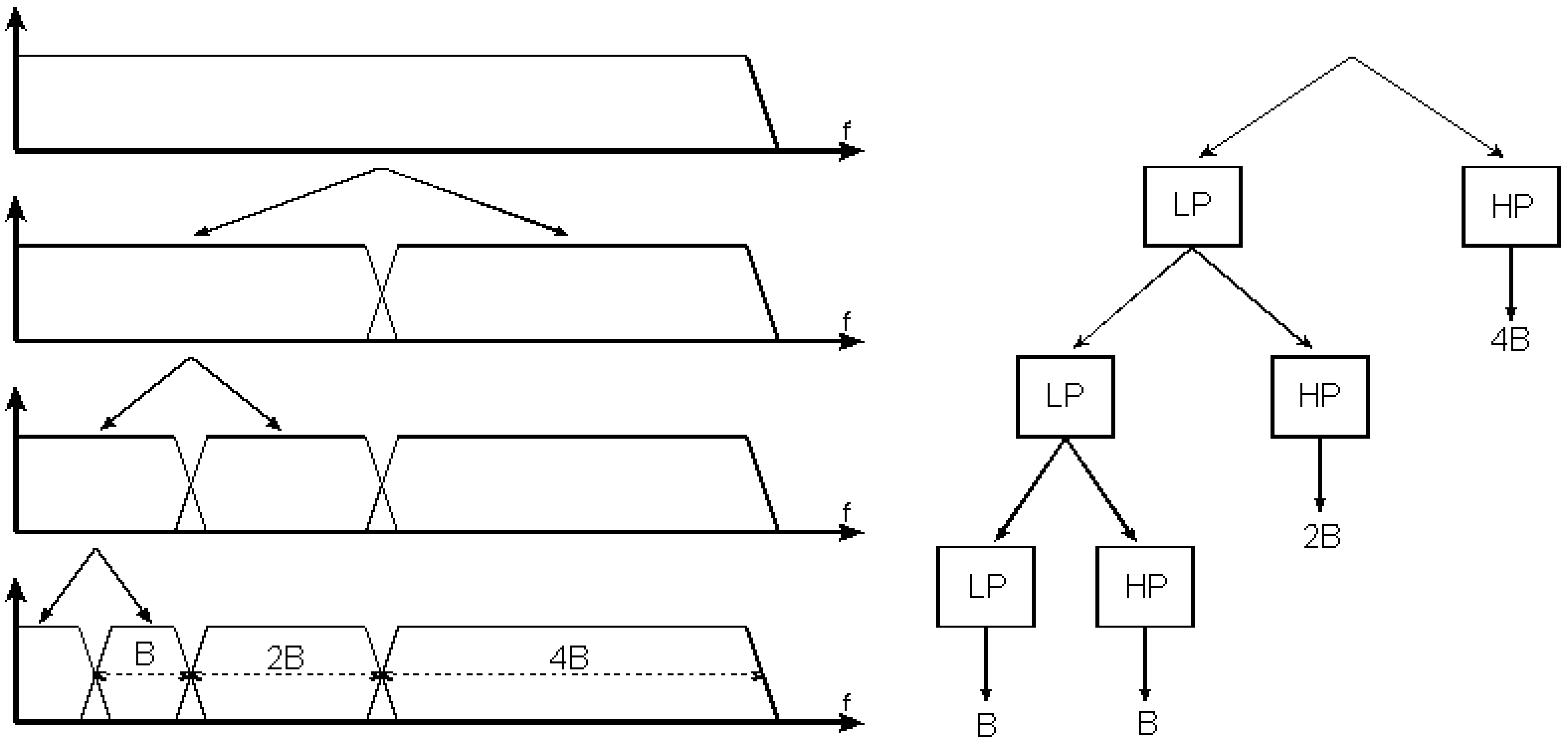
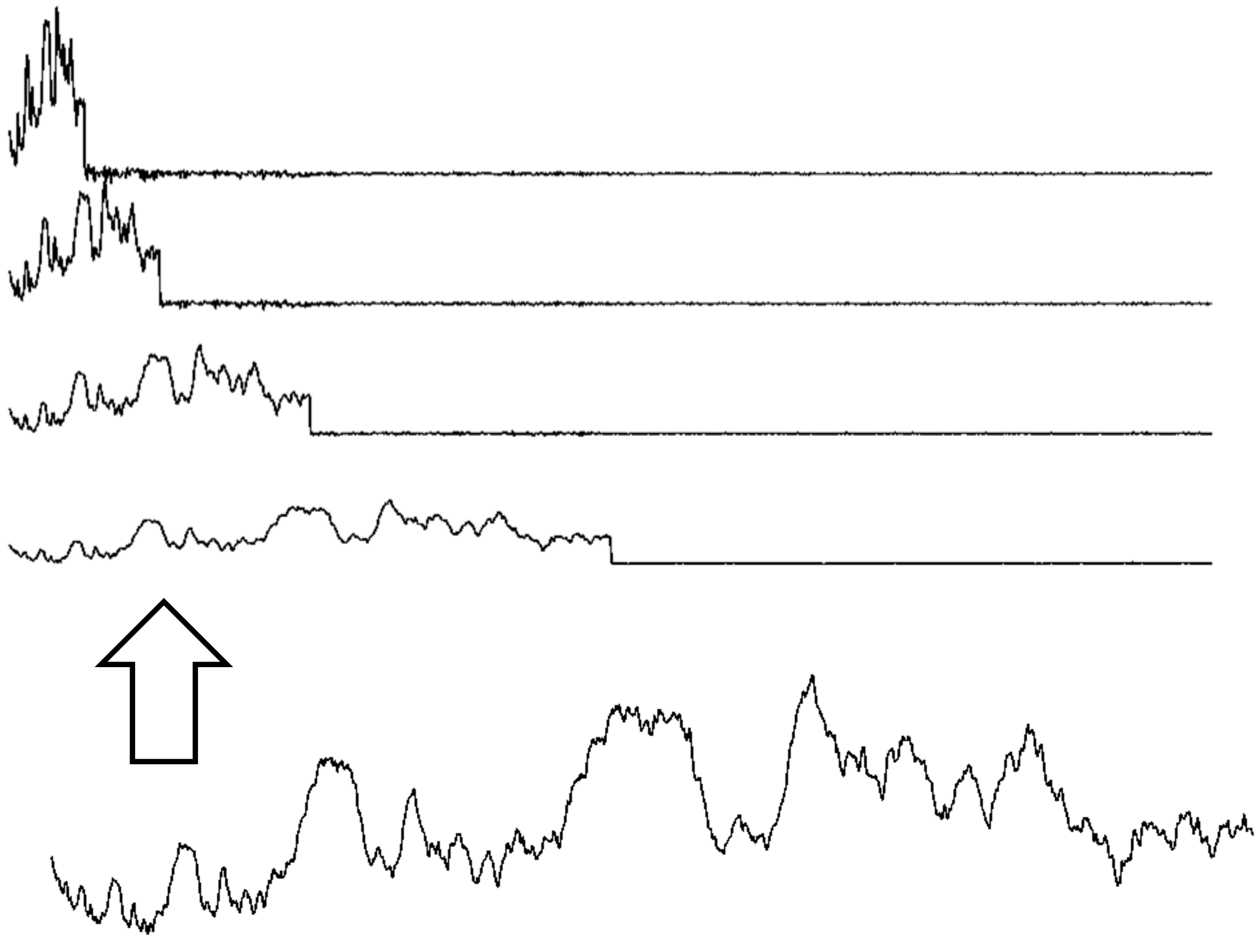


Image from http://www.polyvalens.com/blog/wavelets/theory/

Multi-resolution analysis



Dyadic decomposition

- These points indicate where weights for wavelets with a given translation and scale are evaluated.

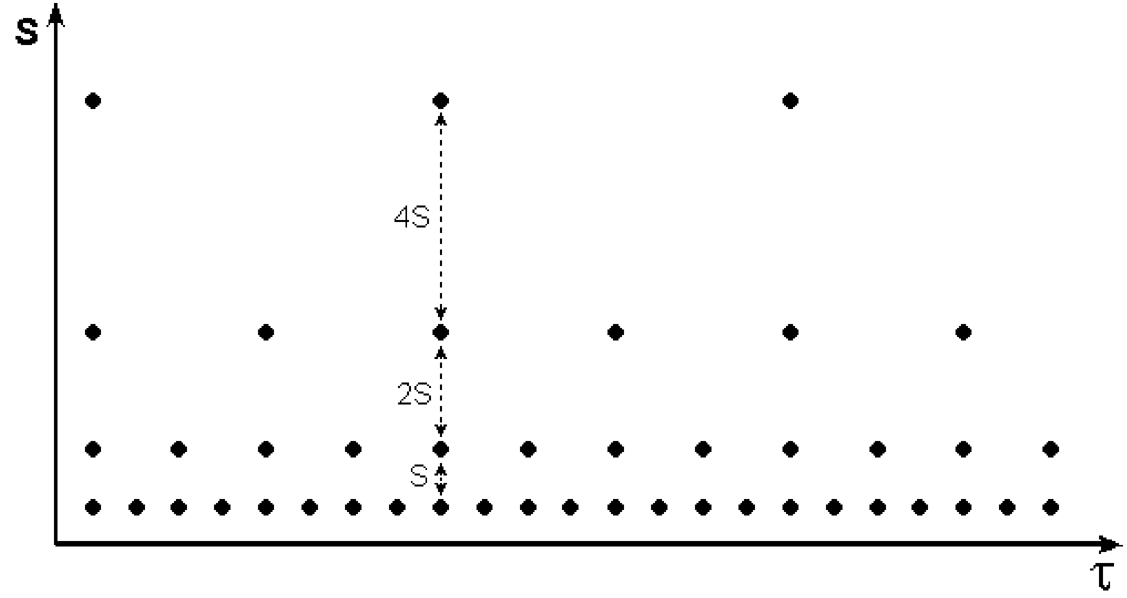
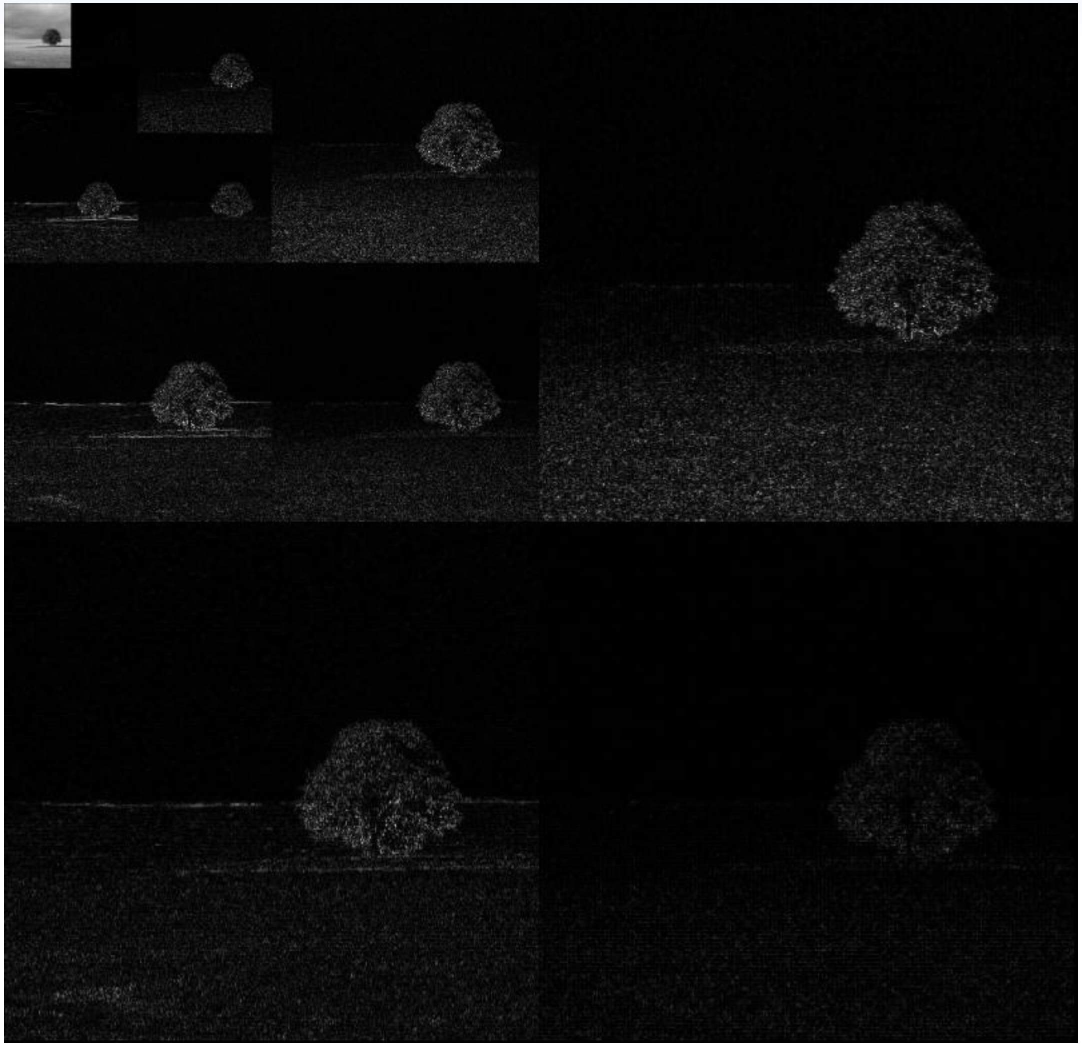


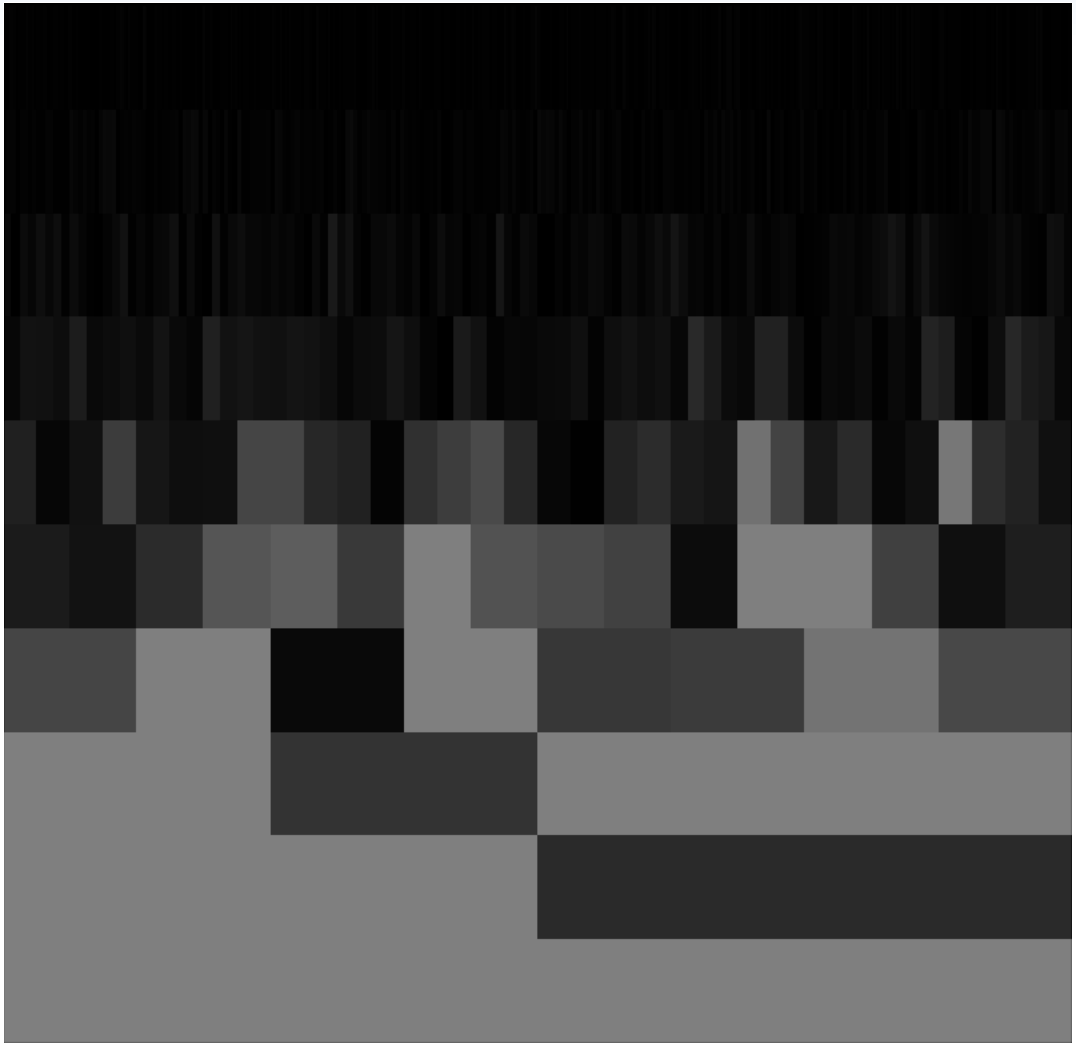
Image from http://www.polyvalens.com/blog/wavelets/theory/

Wavelet decomposition



DWT generated by demo program.

Wavelet decomposition



DWT generated by demo program.

Convolution-based filtering

- Convolve an extended signal with filter masks.

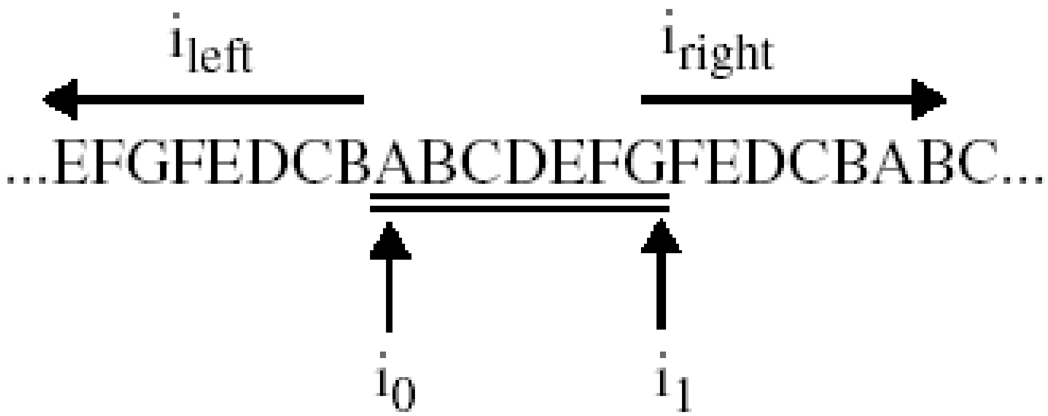


Image from http://www.mathworks.com/help/wavelet/ug/lifting-method-for-constructing-wavelets.html

Lifting-based filtering

- Apply prediction and update operators.

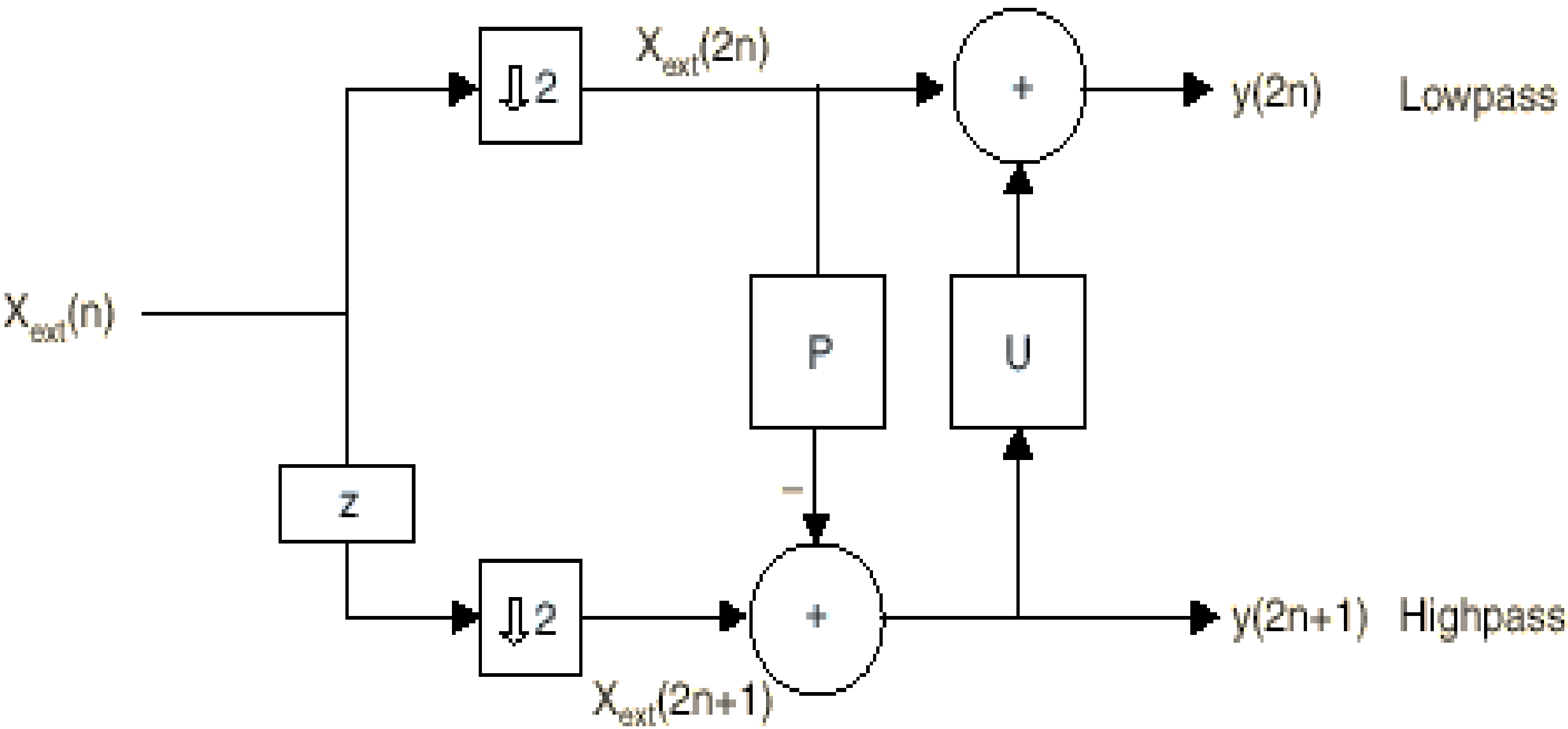
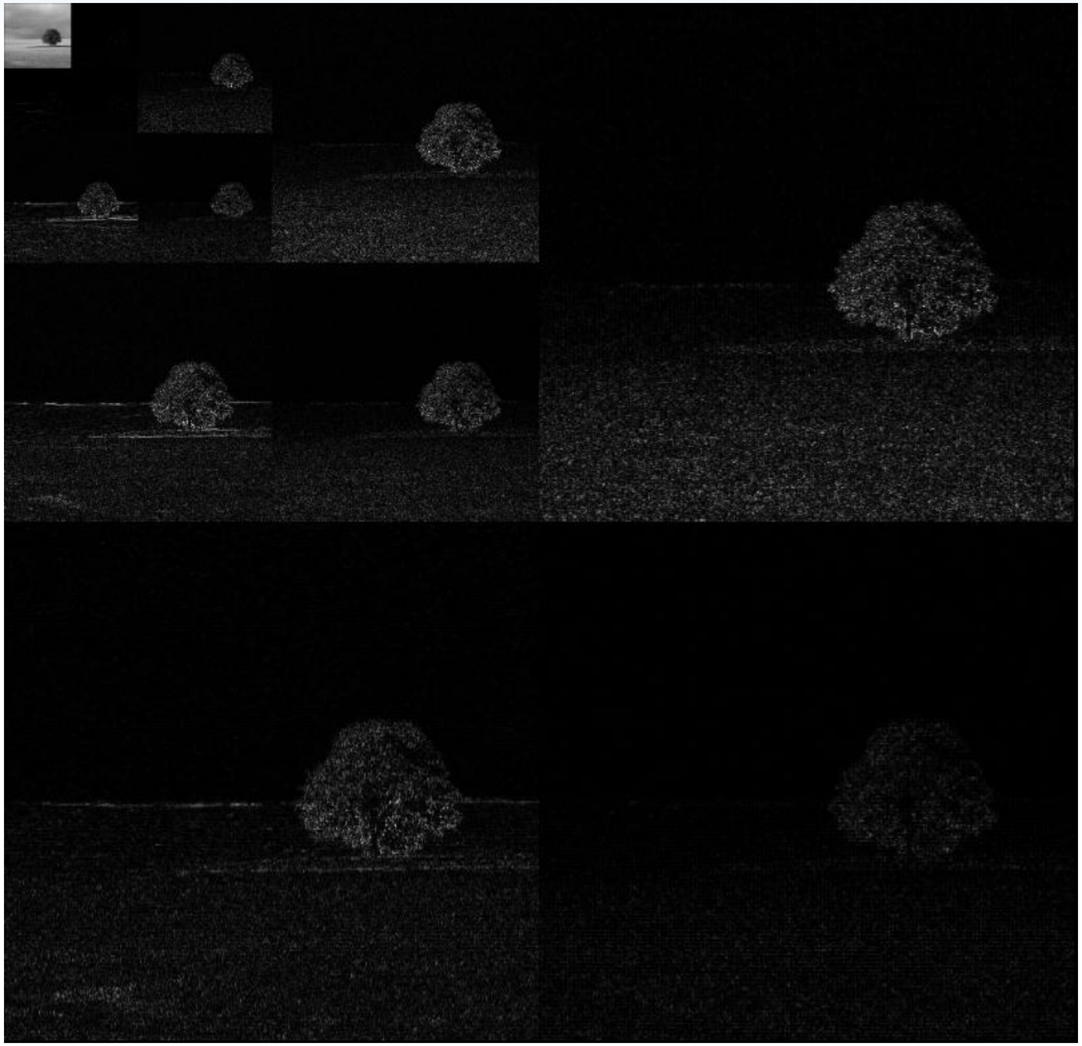


Image from http://www.mathworks.com/help/wavelet/ug/lifting-method-for-constructing-wavelets.html

DWT versus DCT

- DCT is block-based.
  - Blocking artifacts.
  - Not computationally complex – only requires dot products.
- DWT is not block-based.
  - No blocking artifacts.
  - Global detail reduction
  - Filter banks are significantly more computationally complex.

Embedded Zerotree for storage



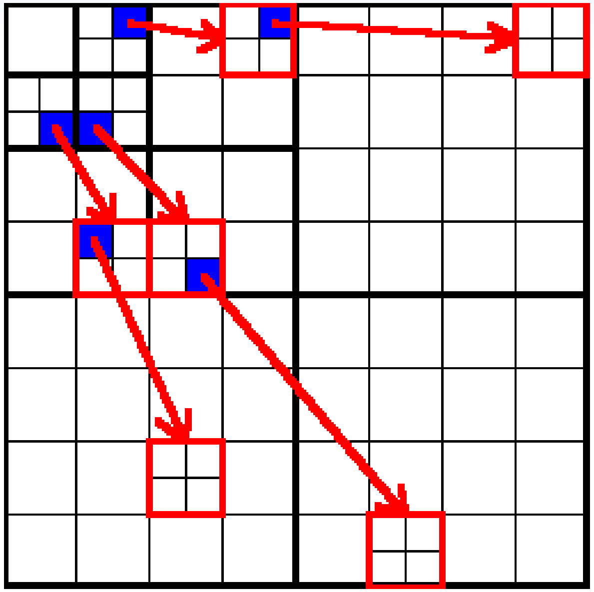


Diagram from **Shapiro**, J. M. *EMBEDDED IMAGE CODING USING ZEROTREES OF WAVELET COEFFICIENTS.* IEEE Transactions on Signal Processing, Vol. 41, No. 12 (1993), p. 3445-3462.

Progressive coding using EZW

- Reconstruct the image from symbols in the order that they are read.
- Without EZW, fully detailed rows would begin to appear from the top.

Image from extras.springer.com/2000/978-3-540-66757-5/kap7/0705203.htm

The most important question

- Which technique should be used?
- It depends on the type of image
- Wavelets are ideal for natural images.
  - Photography
  - Terrain heightmaps
- Wavelets are far from ideal for artificial images.
  - Cartoons
  - Circuit diagrams
  - Sharp discontinuities

The end

- Questions?