
Compressing images for the Internet

Giordano Beretta

Hewlett-Packard Laboratories

<mailto:beretta@hpl.hp.com>

http://www.hpl.hp.com/personal/Giordano_Beretta

© 1998 Hewlett-Packard Company. All rights reserved.



The hotting of the W³

- The W³ was designed for the paradigm that
 - the author decides contents and structure
 - the reader decides the appearance
- Today the W³ is a hot publication medium
 - authors want tight control over the appearance
- Mechanisms¹ like **C**ascading **S**tyle **S**heets (CSS) and **eX**tensible **S**tyle **L**anguage (XSL) are too new and not supported by most deployed browsers

1. <http://www.w3.org/Style/>

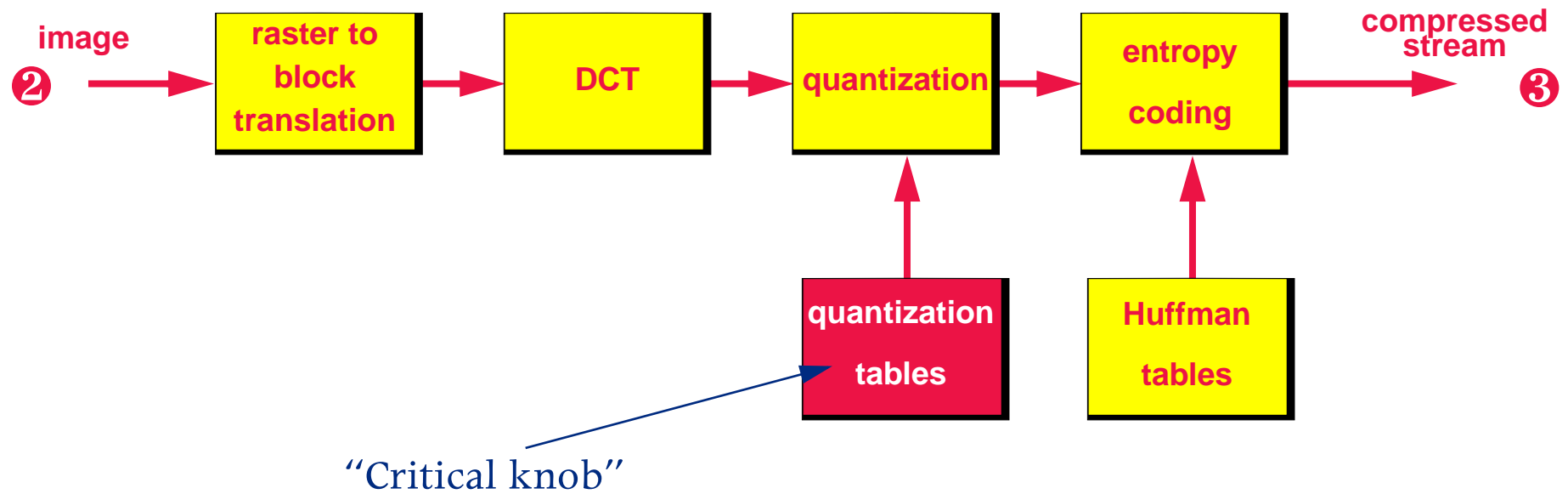
- Two popular image encodings for the W^3
 - GIF for graphics
 - few palettized colors, color accuracy not important, LZW compression
 - JPEG for pictorial images
 - lossy compression, very accurate color communication
- To control appearance, professional publishers use images to communicate text
- This text often has shadows and a textured background
 - GIF has too few colors in the palette
 - representation is not colorimetric
 - JPEG introduces artifacts in graphical images
 - ringing in text
 - blockiness in smooth regions

Examples

- The same image: original, GIF, JPEG



ISO/IEC IS 10918-1 (a.k.a. JPEG)

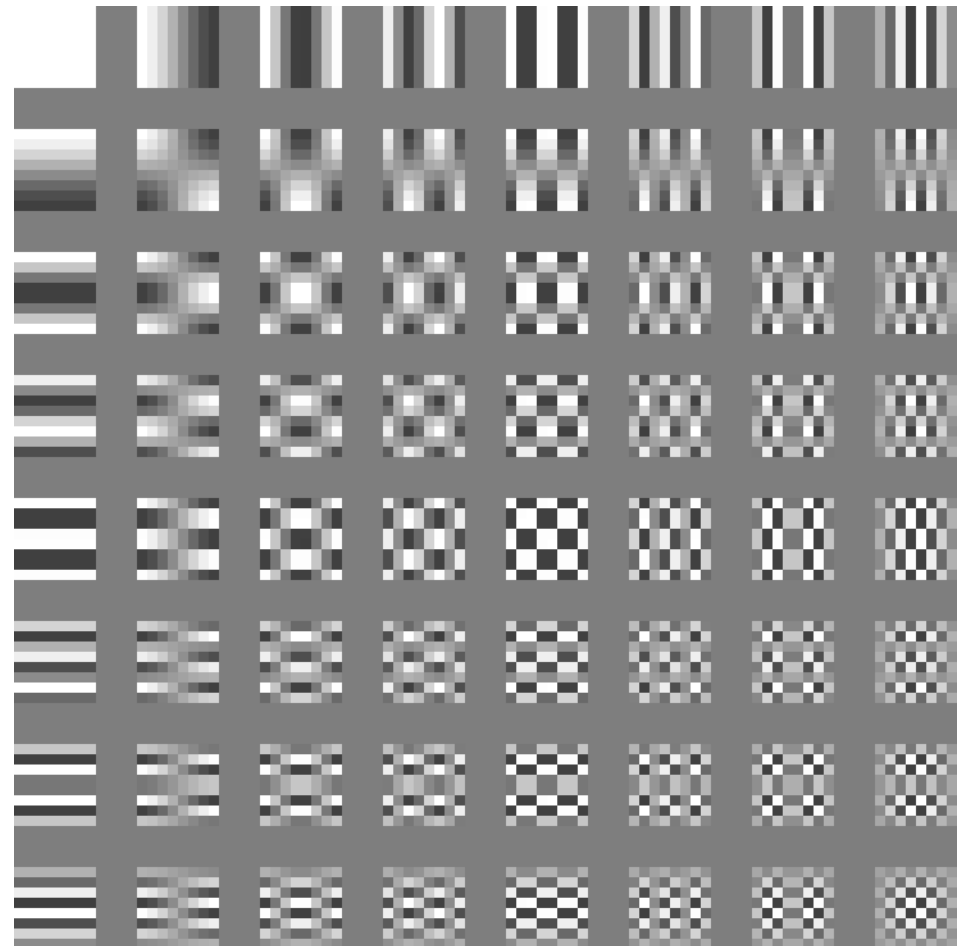


The DCT and its kernels

$$Y(k, l) = \frac{1}{4}C(k)C(l) \sum_{x=0}^7 \sum_{y=0}^7 S(x, y) \cos \frac{(2x+1)k\pi}{16} \cos \frac{(2y+1)l\pi}{16}$$

$$[C_8]_{mn} = k_m \cos \frac{m\left(n + \frac{1}{2}\right)\pi}{8}$$

The 64 kernels of the discrete cosine transform:



Classical approach: the q-factor

The same image compressed with the same parameters except for an increasing q-factor



Bit-rate control techniques. Allocate bandwidth in proportion to the average energy

- Works on some image classes, but is not in the spirit of perceptually lossless compression

Contrast sensitivity techniques. Allocate bandwidth in proportion to the human visual system's contrast sensitivity

- Single-resolution theories are simple to implement but the resulting tables are sub-optimal. The technique is useful if the MTF of a particular device must be incorporated

Multi-resolution techniques. Features implicit in the DCT are loosely analogous to a set of oriented spatial-frequency channels. However, features are not the same as for the human visual system (HVS)

- Empirical method: heuristically guess a good set of DCT starting from a known good table
 - advantage: experts can quickly create a good-enough table
 - disadvantage: requires psychophysical experiments
- Systematic method: iterate towards an optimal table using a model of the HVS
 - advantage: automatically creates excellent tables
 - disadvantage: very hard to implement (Beau Watson, Mike Brill, ...)
- DQT can be used for all images of the same class
 - but: image is created only once, downloaded many times
 - it may be more efficient to compute custom tables for each image
- Combine with device MTF if the output device is fixed and known (use "best case" for a small set of similar devices)

Simplification for design:

- Set to 2^{-16} all symbol probabilities less than 2^{-16}
- Proceed as if no constraints on the codeword lengths
- For each image:
 1. compute probability distribution of all possible symbols for which Huffman code words are needed
 2. design image-dependent custom tables
- Gain approximately 5% additional reduction of file size

image	raw	GIF	PNG	TIFF	JPEG	ratio
drivers	24,975	4,703	11,090	9,698	1,753	1:14
main 1	120,540	27,732	90,860	78,372	11,146	1:11
main 2	88,770	18,307	54,747	46,614	6,724	1:13
main 3	523,392	116,265	394,306	349,082	39,129	1:13
main 4	130,872	16,890	50,013	46,944	8,088	1:16

raw: uncompressed bitmap, no headers

GIF: non-interlaced, adaptive palette, 256 colors

PNG: full color, full α channel, no filtering, non-interlaced

TIFF: full color, with LZW compression

JPEG: baseline, custom tables

- FlashPix image format

- Images stored in resolution pyramid
- Each plane is tiled (64×64 pixels)
- Each tile is compressed with JPEG
- OpenPix protocol
- Requires plug-in

- PNG image format

- pronounced "ping"
- full-color & colorimetric, γ support, in addition to indexed device-dependent color modes like GIF
- unencumbered replacement for GIF (patent-free)
- lossless compression using LZW (like GIF and TIFF)
- not yet widely supported in deployed browsers

