Encoding Color Images for the World Wide Web

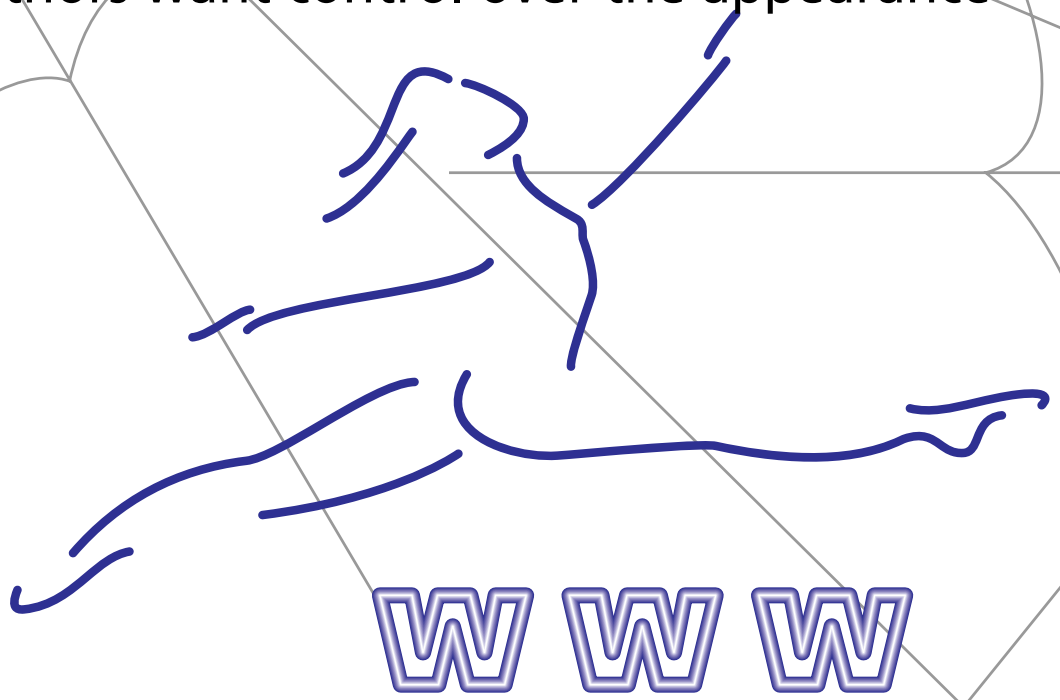
Giordano Beretta, Vasudev Bhaskaran, Konstantinos Konstantinides, Balas Natarajan

Hewlett-Packard Laboratories

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

The hotting of the W^3

- The W^3 was designed for the paradigm that
  - the author decides contents and structure
  - the reader decides the appearance
- Today the W^3 is a hot publication medium
  - authors want control over the appearance



Problem

- Two 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

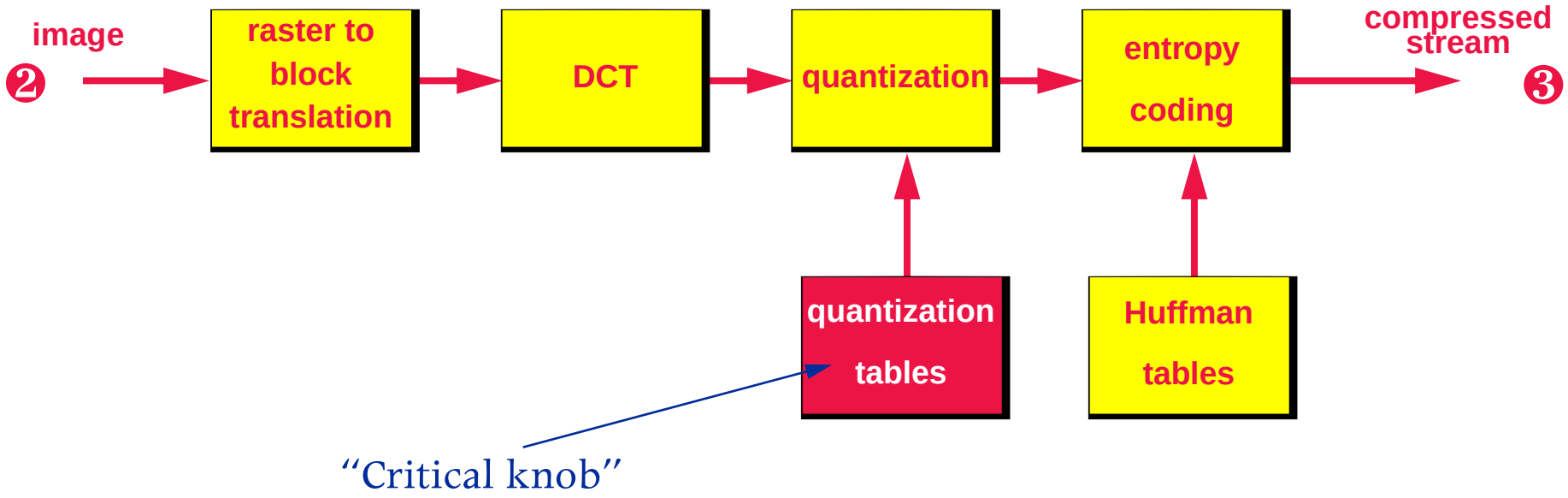
Color in JPEG

- No color space specification
- Baseline JPEG: 4 or less color components
  - Colorimetric color representation is possible
- Full JPEG: 256 or less color components
  - Discrete spectral color representation is possible

Conclusions:

- Unlike GIF, JPEG can be used for accurate color communication
- Find way to solve artifact problem in JPEG

ISO/IEC IS 10918–1 (a.k.a. JPEG)

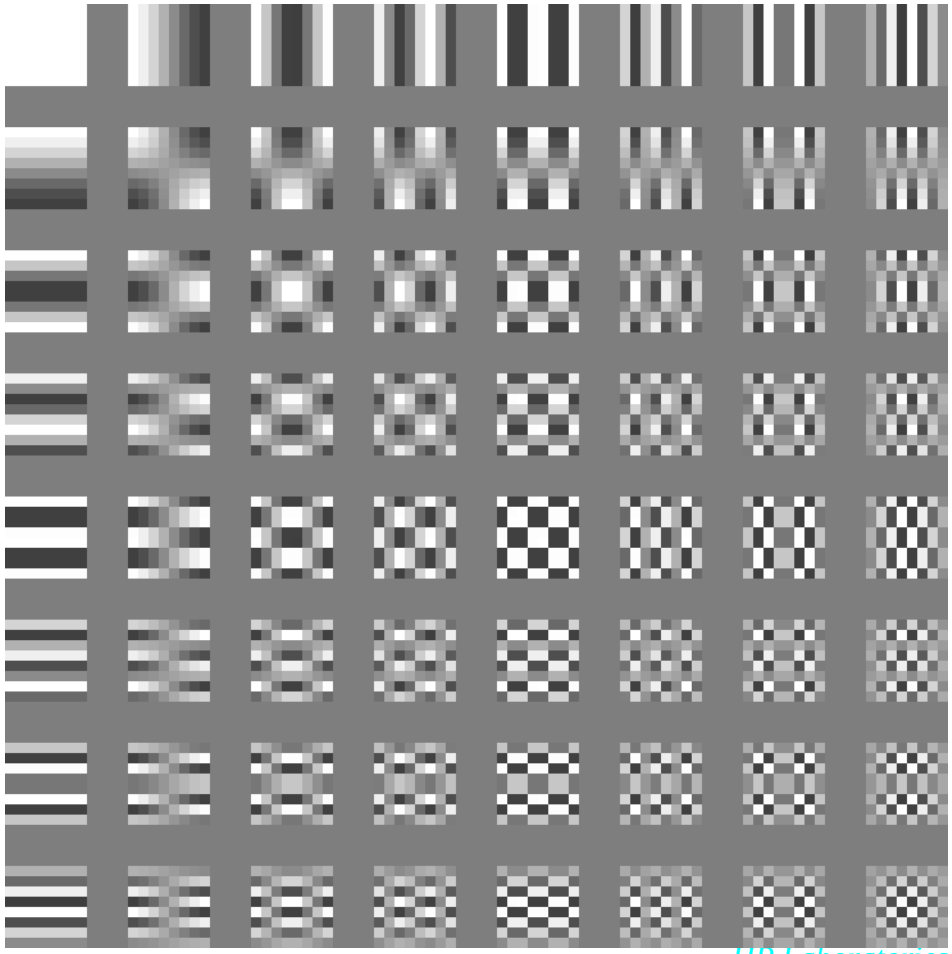


The DCT and its kernels

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

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

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 allocation

DCT transforms 2-dimensional data to a 64-dimensional space

- Each dimension represents a spatial pattern
- For each image or image class:
  - **measure the energy in each of the 64 dimensions**
  - **popular estimator for energy: statistical variance**
  - **allocate bandwidth in proportion to the average energy**
- Better compression rate than with the default tables
- Bit allocation $N_{k,l} = \frac{1}{2} \cdot \log \frac{V_y[k,l]}{D}$ based on variance

$$V_y[k,l] = \text{var}\,(Y[k,l]) = \frac{1}{B} \sum_{i=1}^{B} (Y_i[k,l] - M_y[k,l])^2$$

Strategies

- DQT can be used for all images of the same class
  - text
  - business graphics
  - maps
  - drawings
  - gradients in various directions
  - etc.
- But: image is created only once, downloaded many times
- It is more efficient to compute custom tables for each image: adaptive algorithm

Huffman Tables

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 approx. 5% additional reduction of file size

Results

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, α, no filtering, non-interlaced

TIFF: full color, with LZW compression

JPEG: baseline, adaptive tables

Conclusion

Accurate color communication with short transmission time:

- use colorimetric color representation
- use JPEG image format
- for each image design custom DQT
- for each image/DQT design custom HT

New trend

- FlashPix image format
- Examples: http://image.hp.com/
- Images stored in resolution pyramid
- Each plane is tiled (64×64 pixels)
- Each tile is compressed with JPEG
- Communication protocol allows transfer of individual tiles or groups