Image Processing For Color Facsimile

Giordano Beretta

Hewlett-Packard Laboratories
Imaging Technology Department
1501 Page Mill Road
Palo Alto, CA 94304–1126
http://www.hpl.hp.com/personal/Giordano_Beretta

Joint Work With

- Vasudev Bhaskaran
- Konstantinos Konstantinides
- Daniel T. Lee
- Ho John Lee
- Andrew H. Mutz
- Balas K. Natarajan

Outline

- Background for the standard
- Discrete nature of visual perception
- JPEG data compression
- Optimizing the JPEG compression
- Perceptually lossy compression
- Text sharpening
- Conclusions

Three Breakthroughs

- Digital imaging — compression algorithms
- Hardware cost / performance — SOHO market
- International standard — ITU-T T.42 Addendum

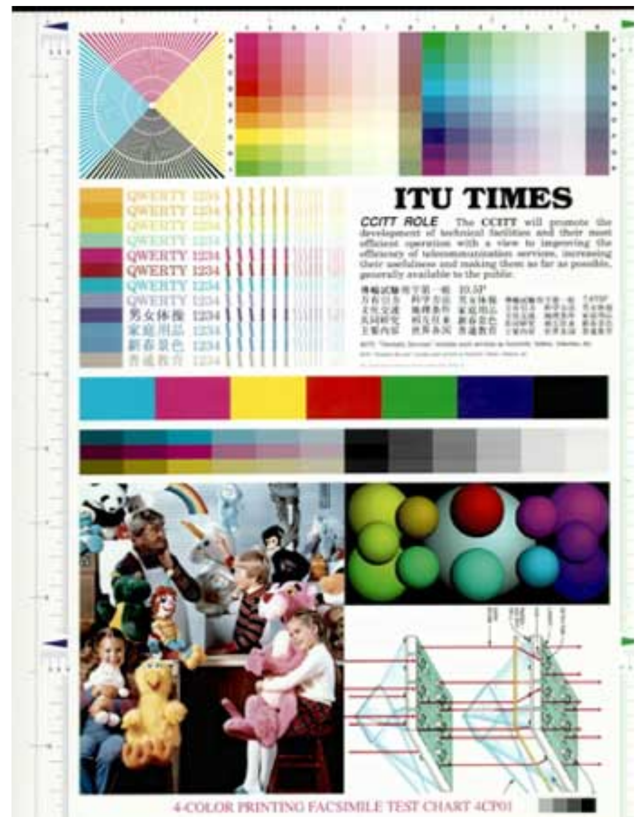
SOHO: Small Office — Home Office

Project Goal

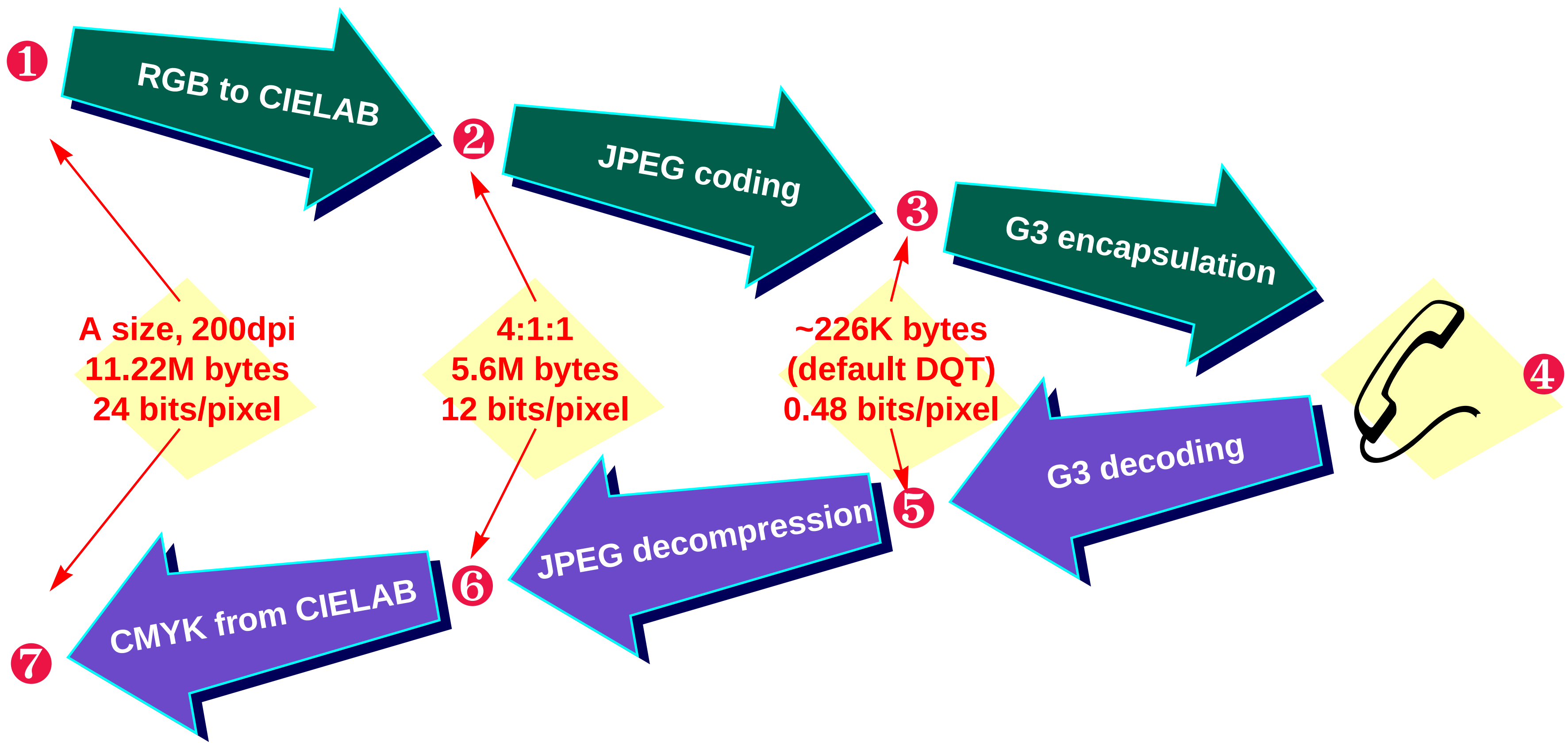
Achieve same transmission time for full color

as for binary black & white

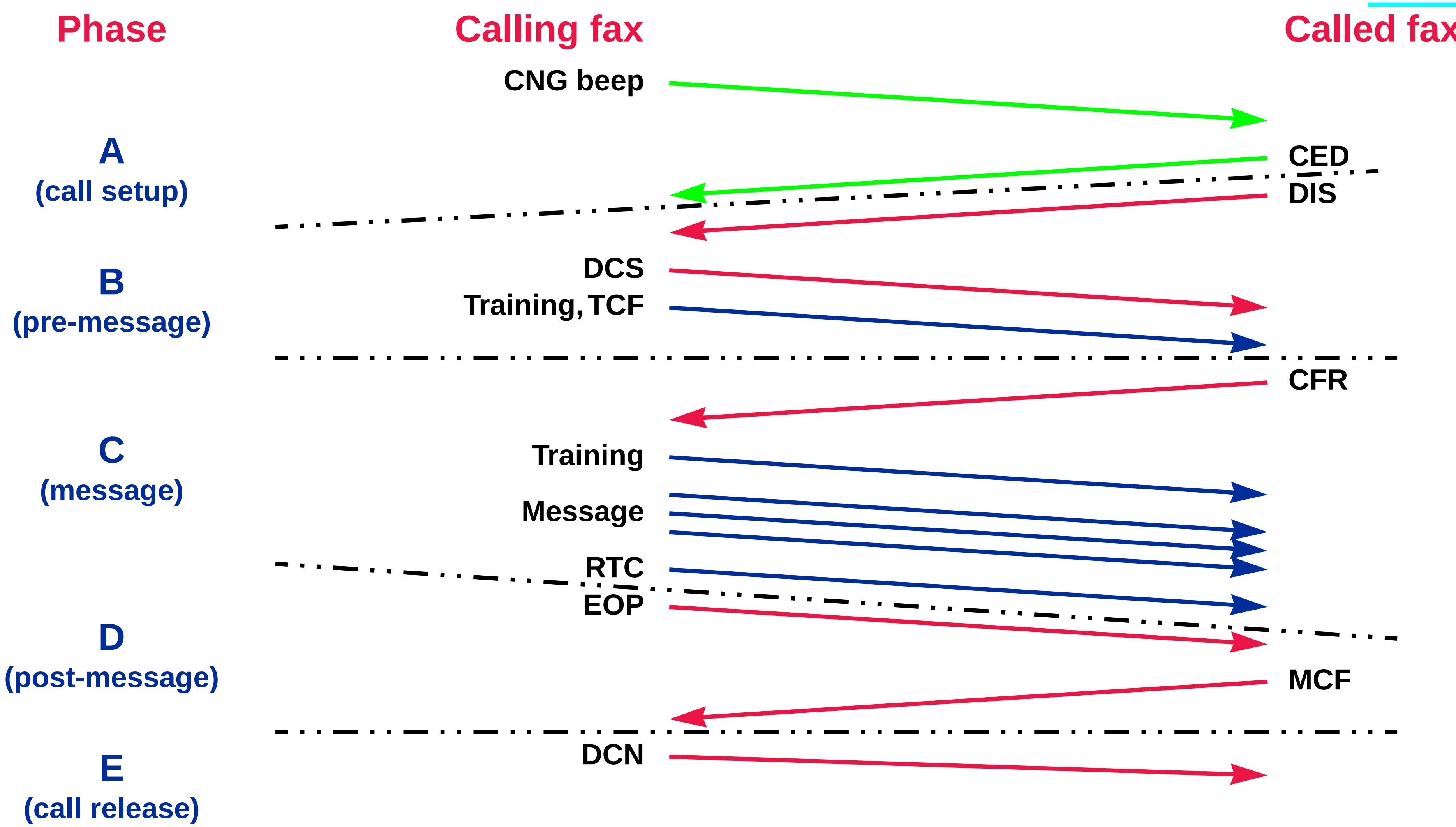
in the case of the 4CP01 test chart



The Color Facsimile Pipeline



Time Phases for T.30 Fax Transmission



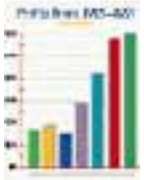
New G3 Entries to DCS Frames

Bit No.	DCS
68	JPEG coding
69	Full color mode
70	Preferred Huffman tables
71	12 bits/pel/component
72	Extend field
73	No subsampling (1:1:1)
74	Custom illuminant (not used)
75	Custom gamut range

Four Categories of Business Images

1. Full color (pictorial, color photographs)

2. Multi-color (color charts & graphs)



3. Bi-color (documents marked up with red ink)

4. Mixed color (combination of 1–3, such as color pages of magazines)

Source: NTT

Color Space Selection

17 spaces considered (Munsell, CMYK, YIQ, CIE colorimetric spaces)

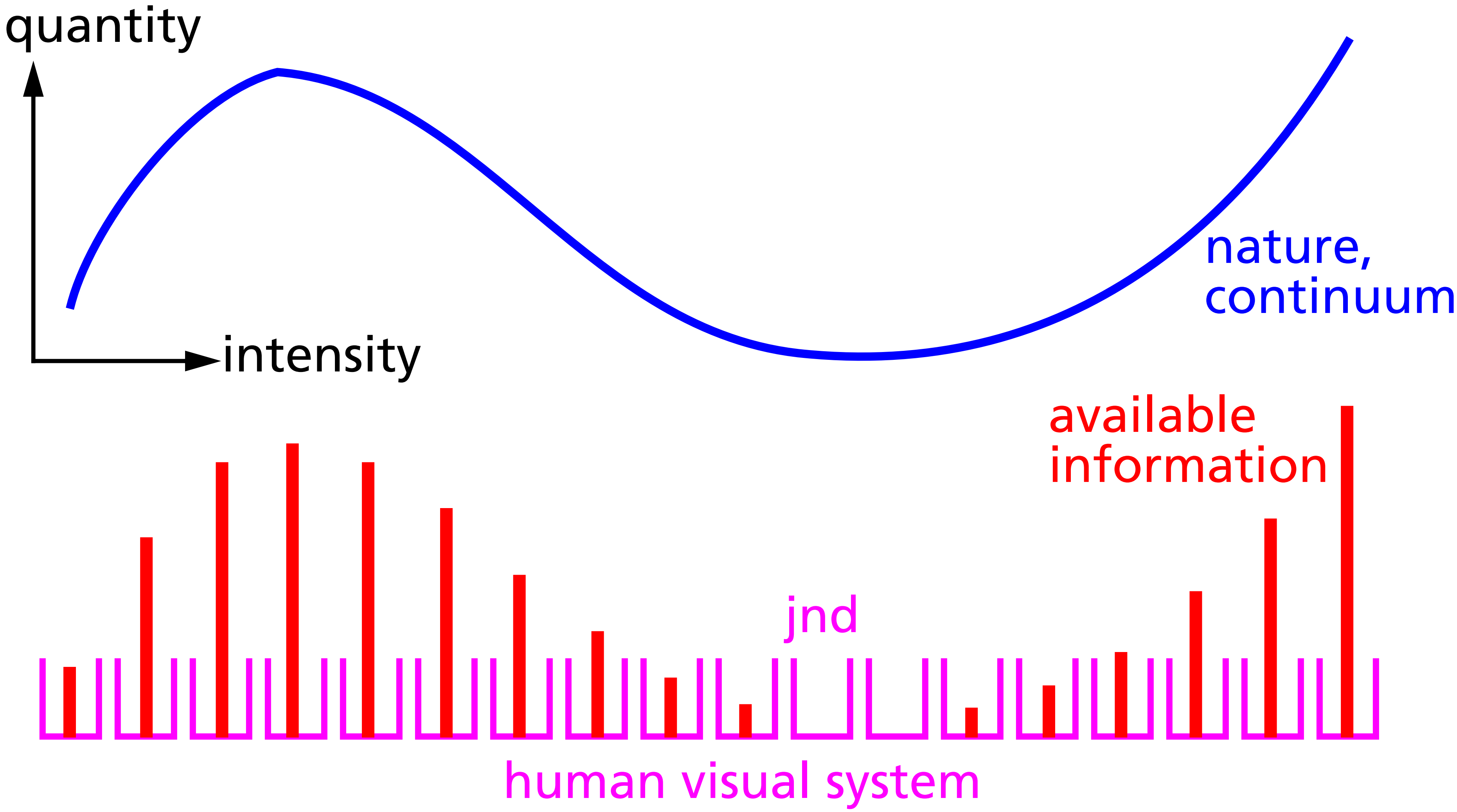
Evaluation criteria (source: Fuji-Xerox):

- ability to represent all colors
- numerical complexity
- device independence
- quantization error under compression
- compatibility with compression algorithms
- color stability with white point change
- …

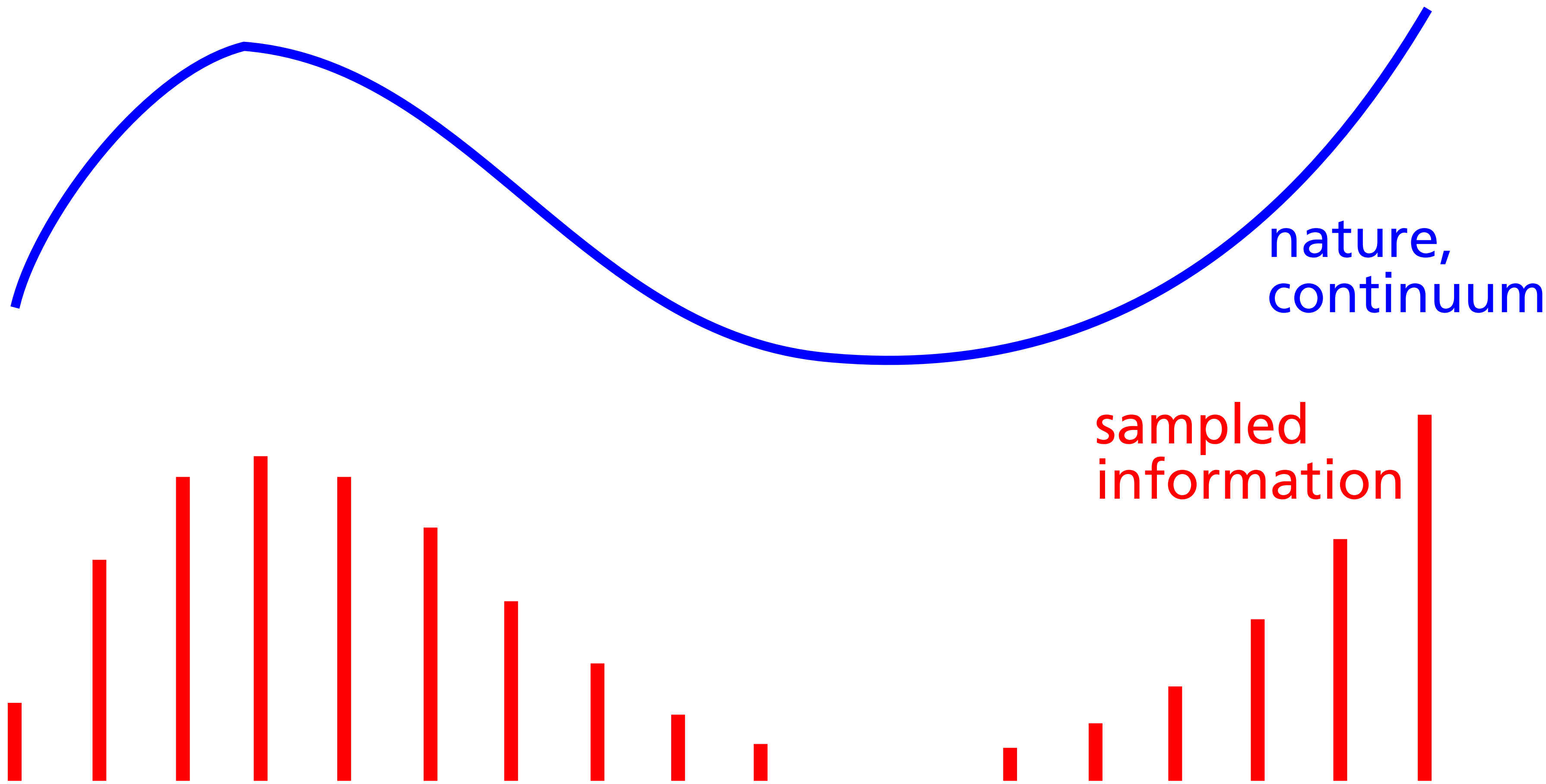
Perception is Discrete

- When a stimulus is changed just by a small amount, an observer will not notice a difference
- Psychophysical experiments: threshold value of where an observer can detect a difference in stimuli
- jnd — just noticeable difference

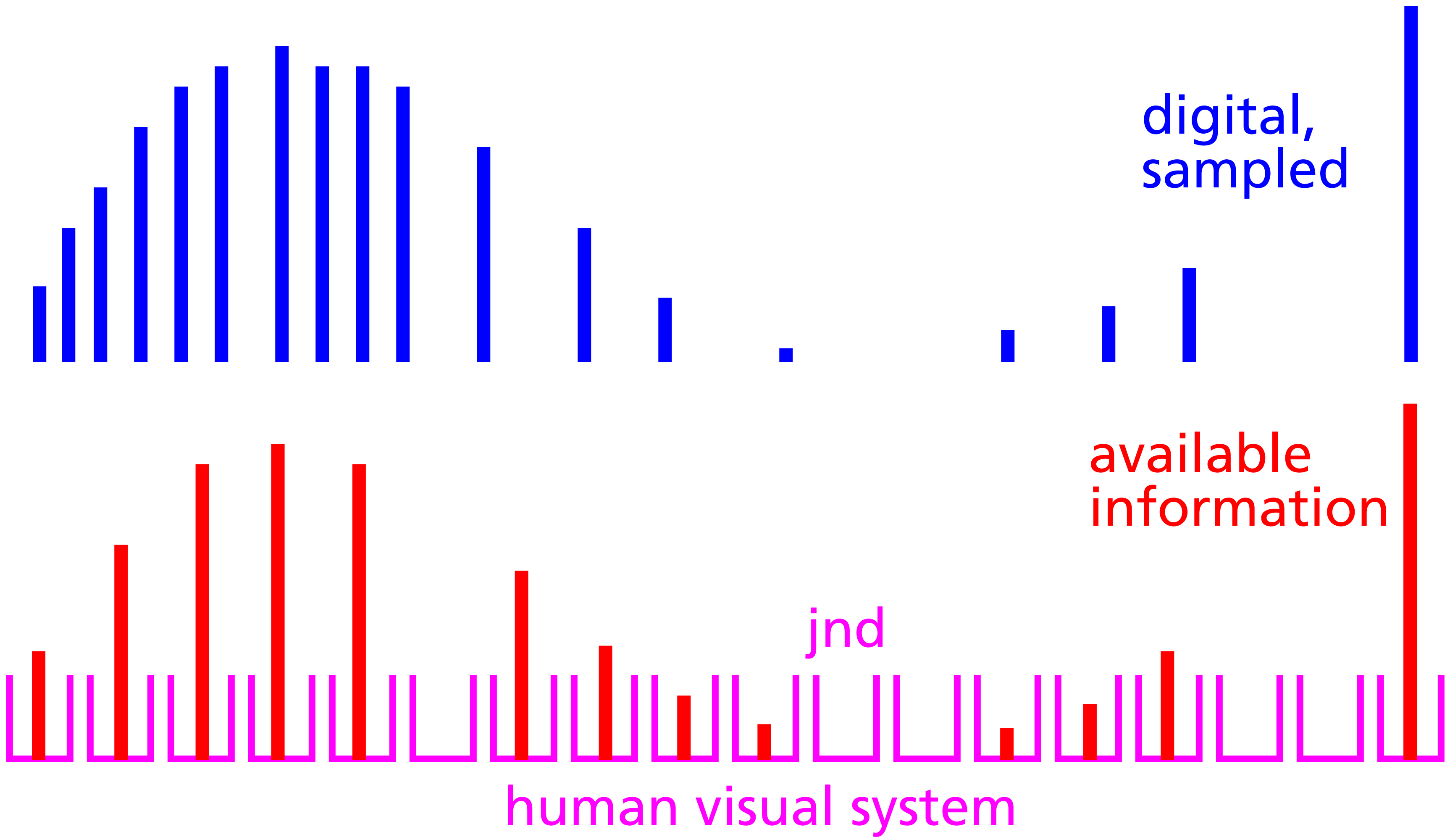
Model for Perception Discretization



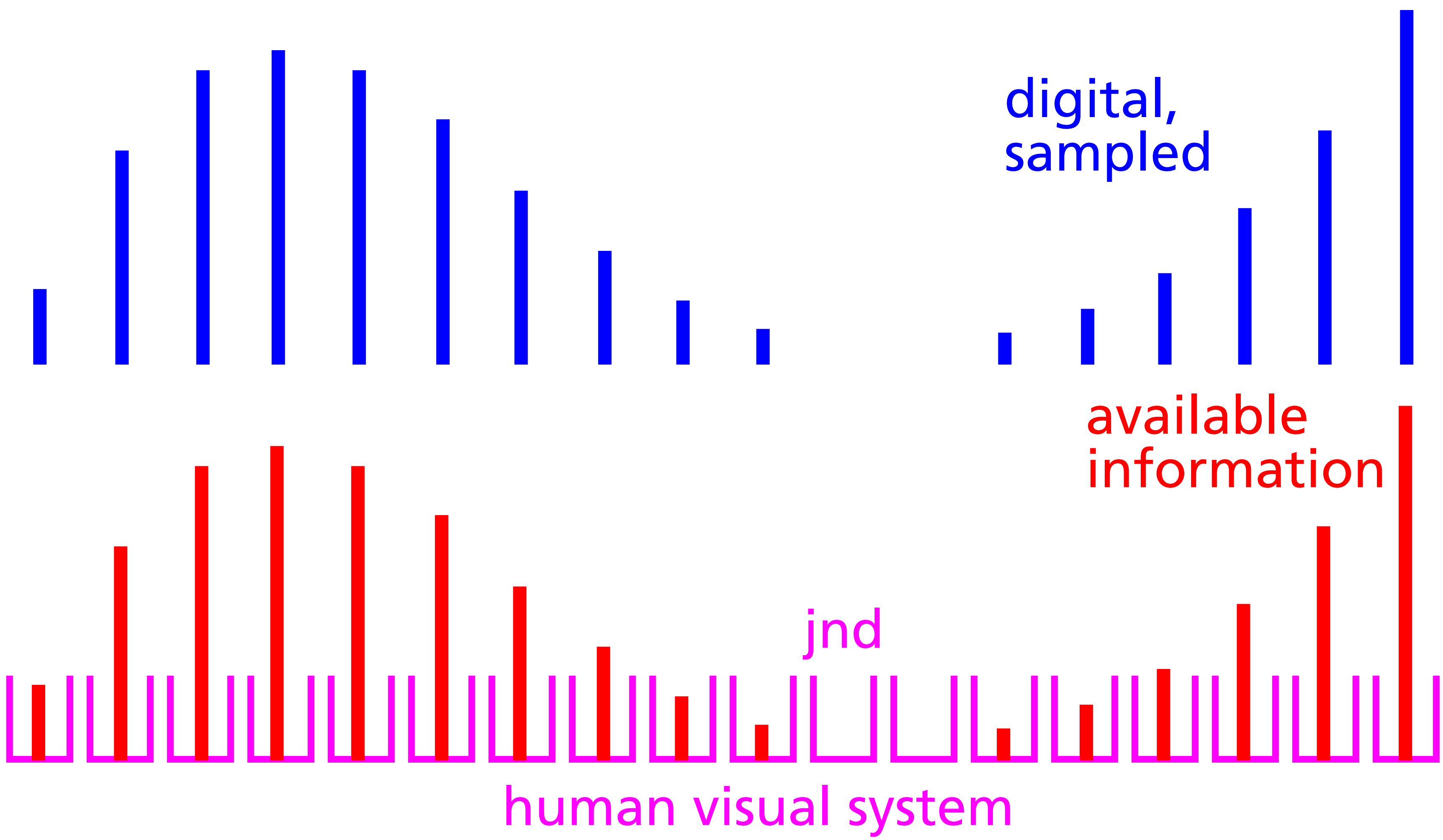
Digital System: Sampling



Poor Sampling

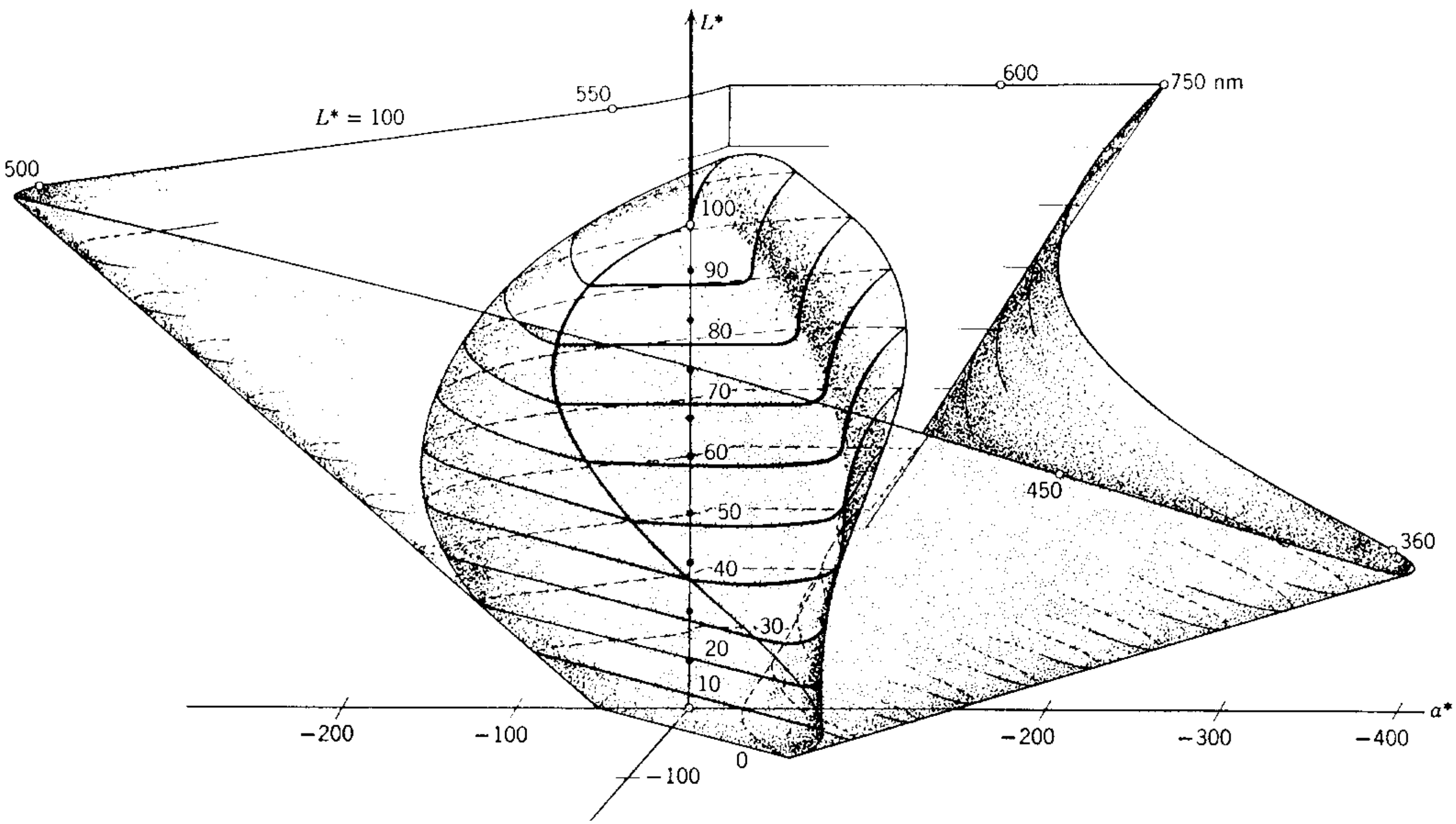


Good Sampling



Result of the Color Space Evaluation

- 1931 Standard Colorimetric Observer
- CIE Standard Illuminant D_{50}
- CIELAB color space

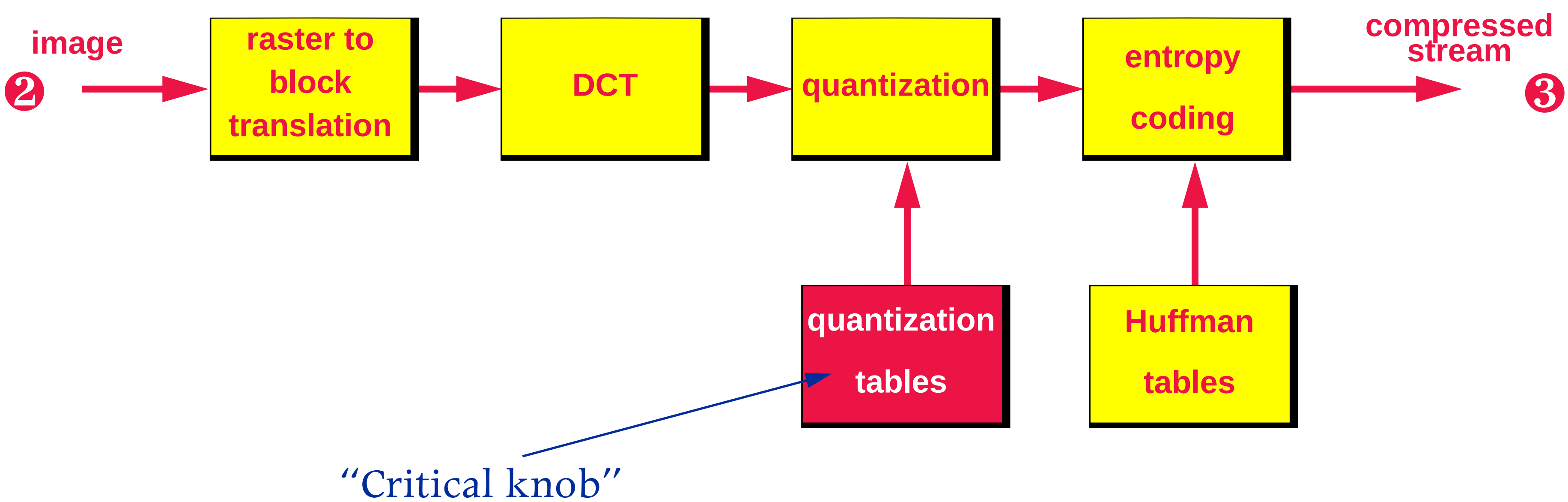


The CIELAB Color Space

- Based on the CIE 1931 Standard Colorimetric Observer: device independent
- Based on Munsell color tree, von Kries adaptation, CIE XYZ color space, and power law compression (n = 1/3): good perceptual uniformity
- Easy to compute compared to other uniform spaces
- Widely used in printing industry
- Can be read directly with measurement instruments
- Design issue: choice of the best range for the chromatic channels $a*$ and $b*$

Data Compression

ISO/IEC IS 10918–1 (a.k.a. JPEG)

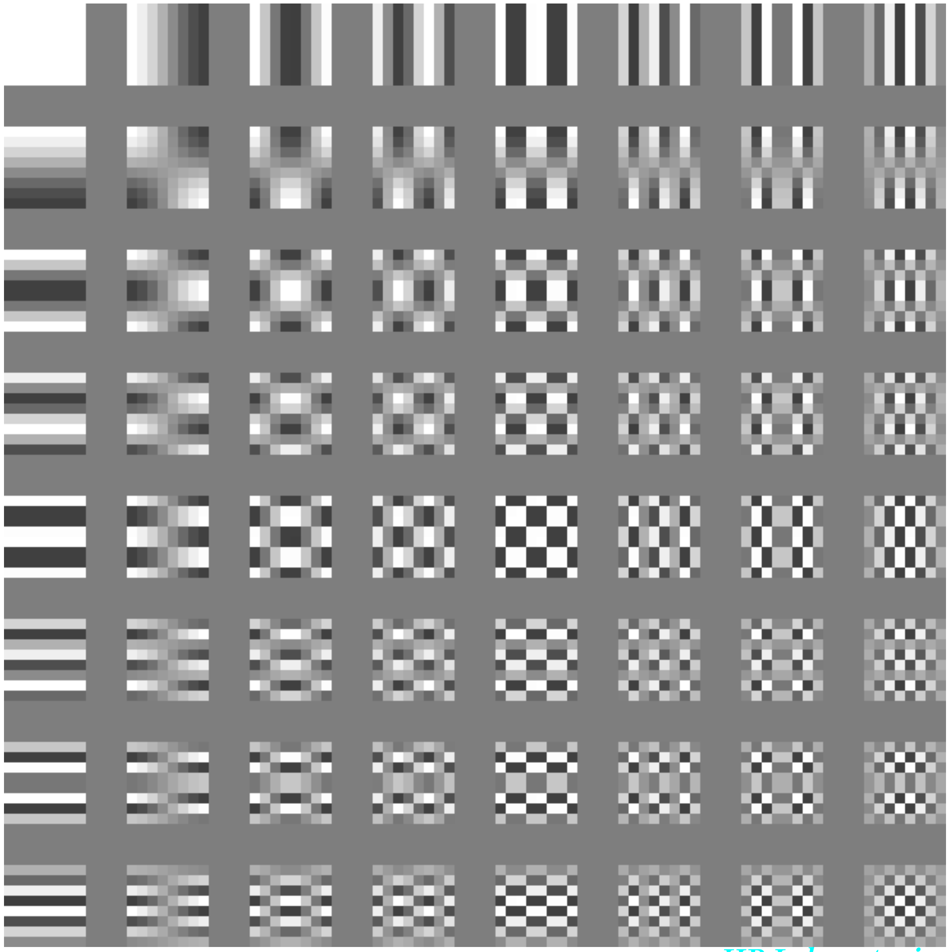


The DCT and its Kernels

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

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

The 64 kernels of the discrete cosine transform:



The JPEG Compression

Quantization tables (DQT) are the key parameter

- Do not quantize where it can be seen
- Default parameters are for images on CRT displays
  - **spatial information in text is different**
  - **printers have a much higher resolutions than CRTs**
- The three worlds of spatial information:
  1. Physical: energy in the signal
  2. Perceptual: sensitivity of the visual system
  3. Semantic: cognitive mechanisms

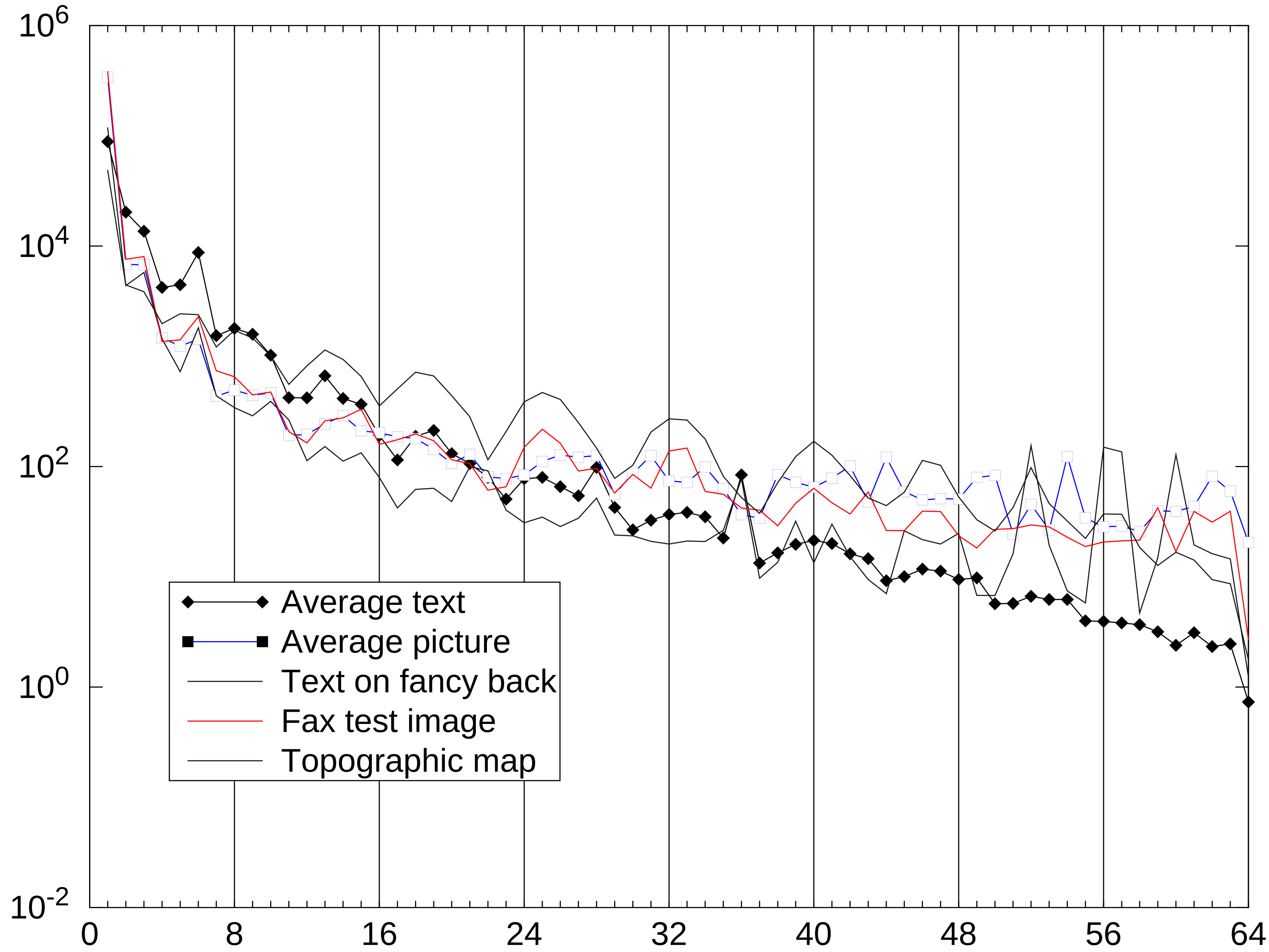
Last step: design the Huffman tables

Physical World

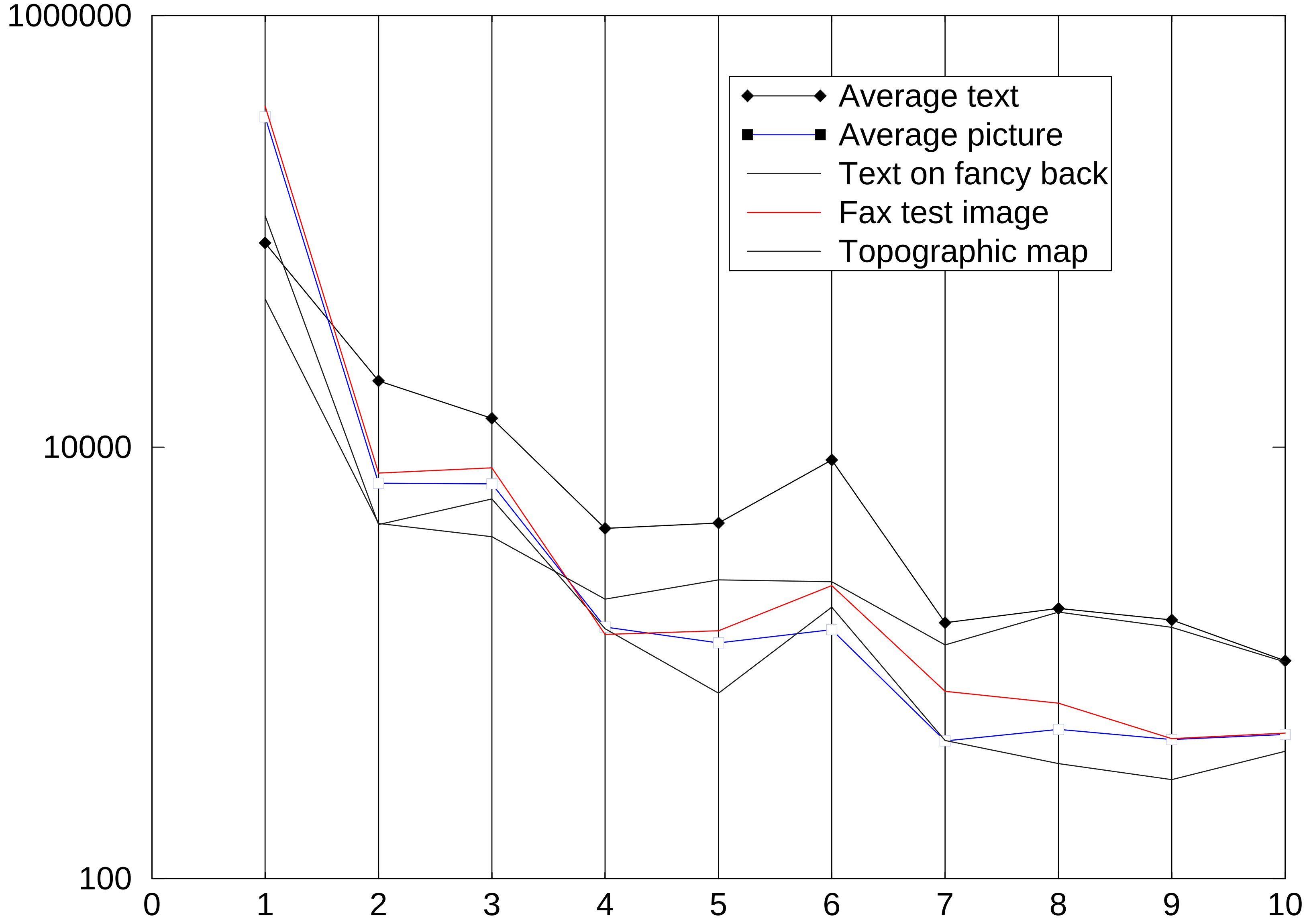
DCT transforms 2-dimensional data to a 64-dimensional space

- Each dimension represents a spatial pattern
- For a number of typical images:
  - **measure the energy in each of the 64 dimensions**
  - **popular estimator for energy: statistical variance**
  - **average over the images in the test set**
  - **allocate bandwidth in proportion to the average energy**
- For text documents with images, a much better compression rate is achieved for a given image quality, than with the default tables

L* Energy in Text vs. Pictorial Images



L* Energy in Text vs. Pictorial Images



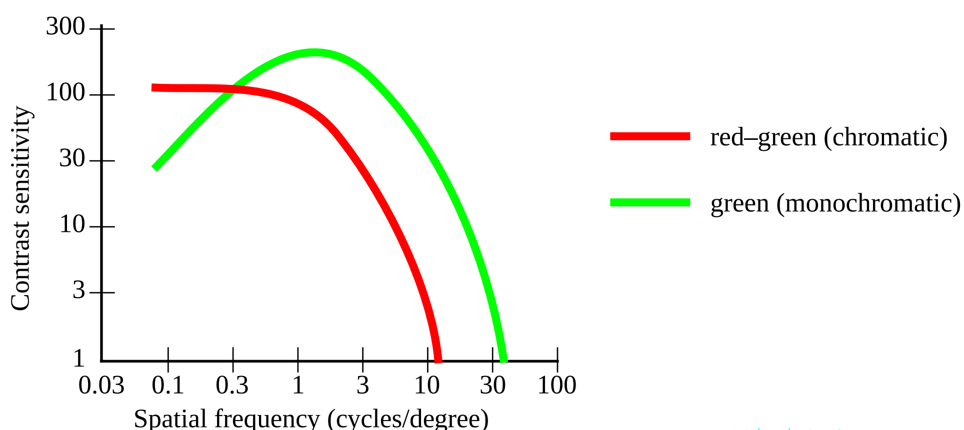
Traditional Bit Allocation

- Bit allocation $N_{k,l} = \frac{1}{2} \cdot \log \frac{V_y[k,l]}{D}$ based on variance

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

To improve bits-per-pixel rate:

1. Brute force: uniform *q*-factor
2. Perceptual: increase DQT elements based on HVS



Perceptual World — Simple Method

Image quality depends on the contrast visible at a given spatial resolution: contrast sensitivity function (CSF)

- Discard spatial information above the CSF: it cannot be seen anyway
- Standard method: weigh the DQT elements by the CSF
- Printer resolution is 600 dpi, same as visual system
  - **improved compression of images**
  - **for text pattern recognition may be more important than CSF**
- Not necessarily a good model for what happens in the visual system

Compression Ratio

Image (10.7M B)	**Default DQT**	**Custom DQT**
Real estate flier with photo	52:1 (211K B)	82:1 (134K B)
Book page with photos and text	53:1 (207K B)	63:1 (174K B)
4CP01 test chart	47:1 (233K B)	63:1 (174K B)

Perceptual World — Complex Method

When two structures (textures) are present in an image, one structure may hide the other

- Principle of visual masking
- When a new structure is added to an existing structure, the new structure may mask the old structure or vice-versa
- Quantization noise is a structure that is added to the image
- Noise that is masked by the image is perfectly acceptable; this allows for higher compression ratios
- This method is used mostly for image-dependent compression (adaptive)

Results

1. If the method is applied to several images of the same type, there is little variation in the obtained DQTs

- Hence, in the case of color facsimile we can use one DQT for all images

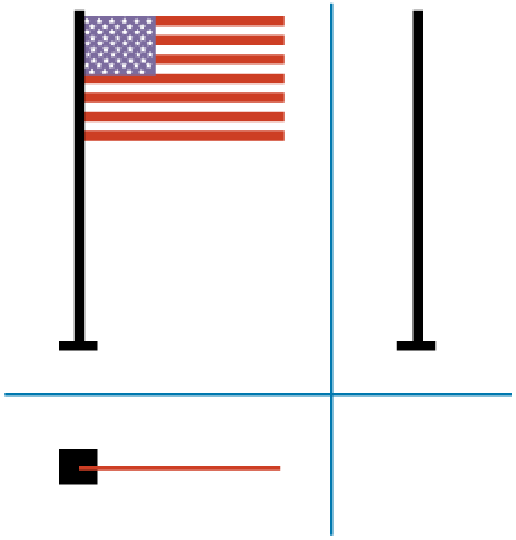
2. Adjust each DQT element to reach the threshold
3. The iterative method converges faster when the previous steps are performed
4. Lossy compression: any jnd value can be targetted

Semantic World

Reading performance of text is the speed at which text can be read without errors

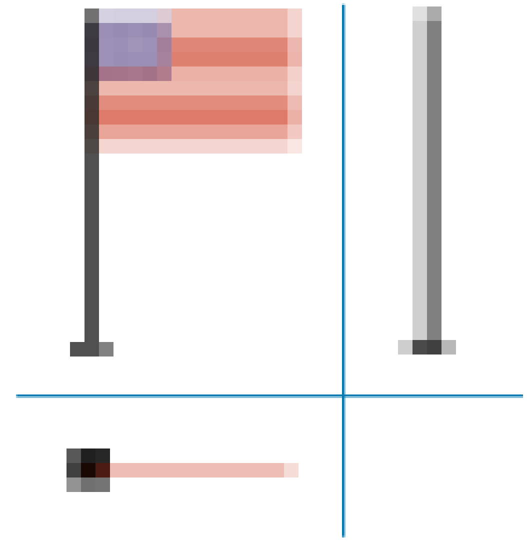
- When compressing, discard information that does not impact reading performance
- Identify the parts of characters in fonts that affect reading performance
- Discard prevalently spatial information not related to these character parts

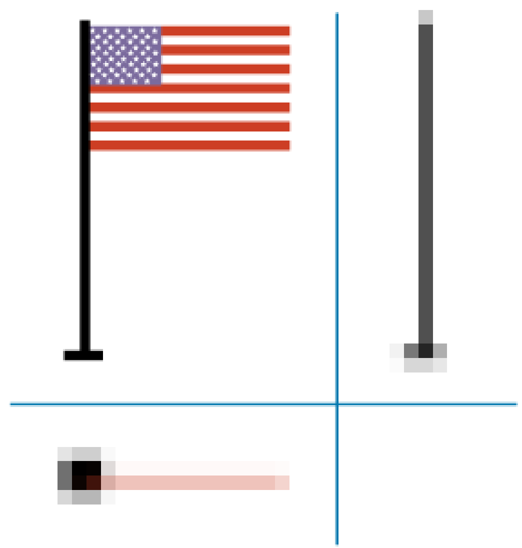
Visual Impact of Quantization Depends on Image and Sub-Space



< fine

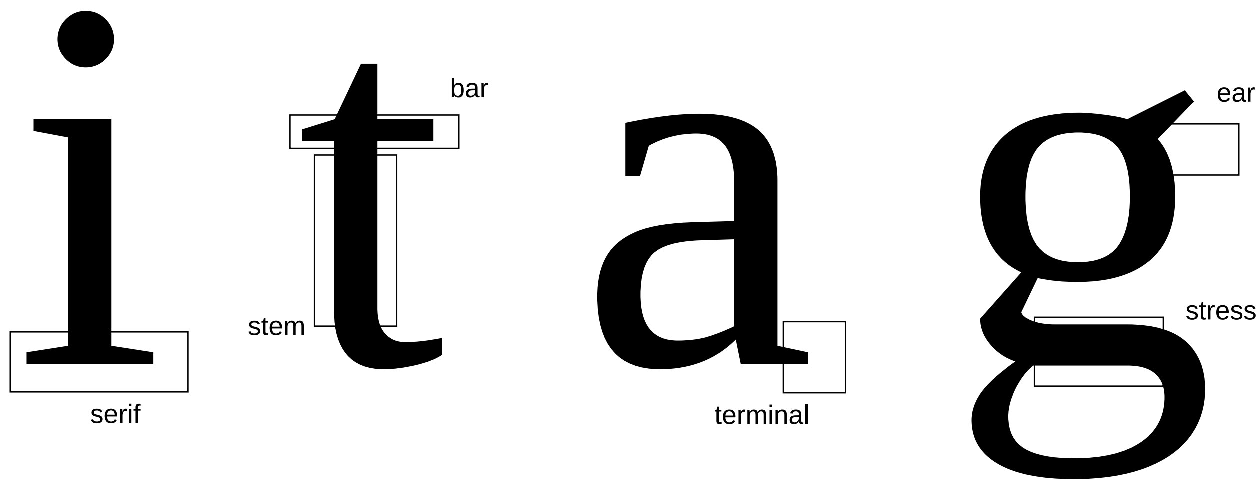
coarse >





< mixed

Some Typeface Parts in Times Roman



7 point	9 point	11 point	12 point	16 point
1.12 mm	1.44 mm	1.76 mm	1.95 mm	2.58 mm

X-Height in the Times New Roman typeface

Critical Feature Sizes in Color Facsimile

Number of pixels	200 dpi	300 dpi	400 dpi
1	0.127 mm	0.085 mm	0.063 mm
2	0.254 mm	0.169 mm	0.127 mm
3	0.381 mm	0.254 mm	0.191 mm
4	0.508 mm	0.339 mm	0.254 mm
5	0.635 mm	0.423 mm	0.318 mm
6	0.762 mm	0.508 mm	0.381 mm
7	0.889 mm	0.593 mm	0.445 mm
8	1.016 mm	0.677 mm	0.508 mm

Red: Limit for visual acuity in the luminance channel
Green: Limit for visual acuity in the chrominance channels
Blue: Peak contrast sensitivity in the luminance channel
Pink: typical character part sizes (11 point Times New Roman)

Modified Bit Allocation Equation

- Weight *w* based on visual quality: $N_{k,\,l} = \frac{1}{2} \cdot \log\left(w[k,l]\frac{V_y[k,l]}{D}\right)$
- Separate provisions for text, graphics, and pictorial images (smoothing)
- Detail of spatial frequency in image data, ind. of phase angle, resides almost totally within 3 coefficients in the transform domain when a DCT is applied.
- Rudimentary example of a weight table:

1	1	1	3/4	1	3/4	1	1
1	3/4	1/2	1/4	1/2	1/4	1/2	1/2
1	1/2	1/4	1/8	1/4	1/8	1/4	1/4
3/4	1/4	1/8	1/8	1/8	1/8	1/8	1/8
1	1/2	1/4	1/8	1/8	1/8	1/8	1/8
3/4	1/4	1/8	1/8	1/8	1/8	1/8	1/8
1	1/2	1/4	1/8	1/8	1/8	1/8	1/8
1	1/2	1/4	1/8	1/8	1/8	1/8	1/8

Preliminary Experimental Results

- Compress two typical 300 dpi images
- Comparable visual quality
- *q* factor 200; use K.2 for the chrominance channels

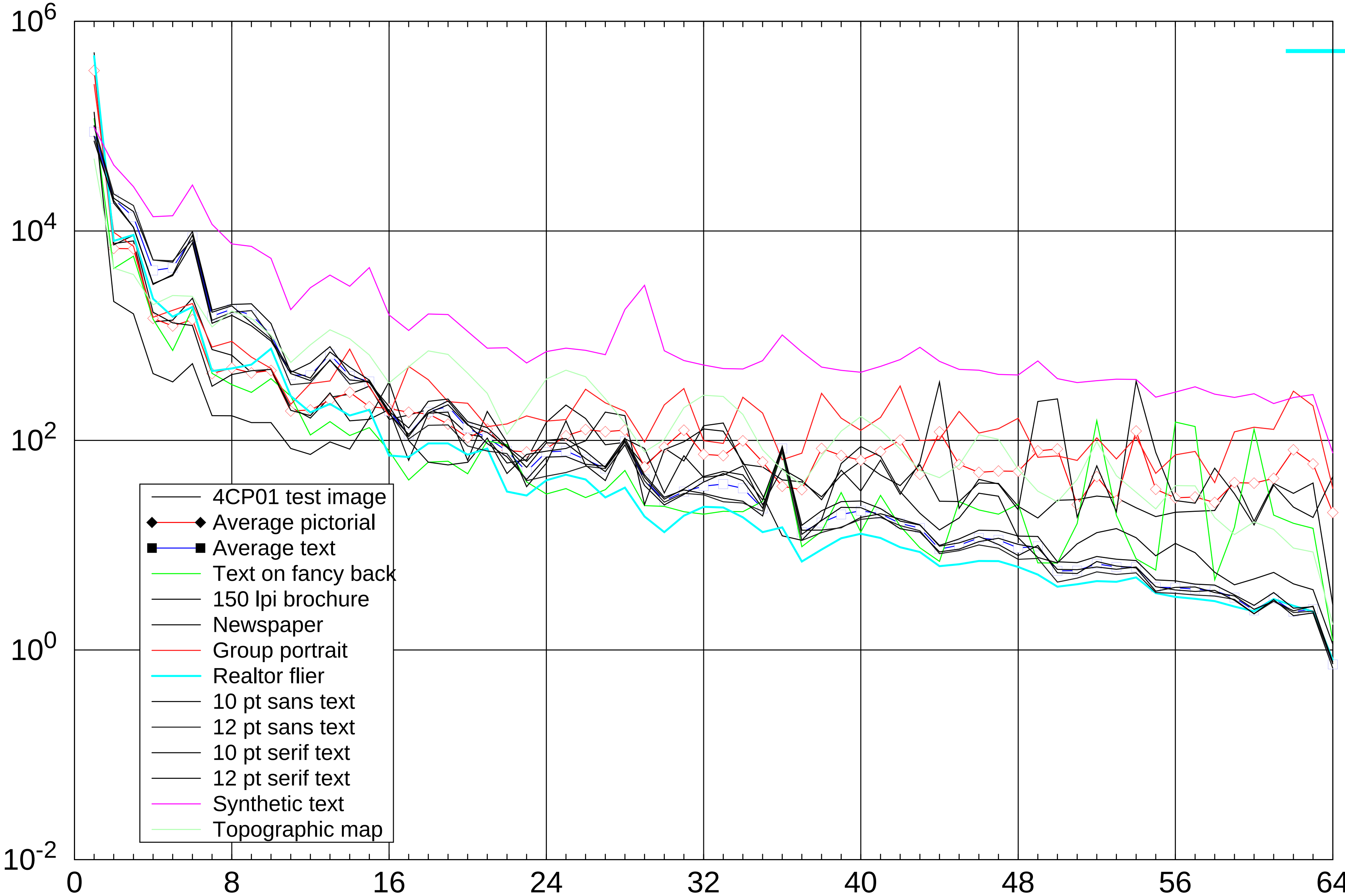
	Text		4CP01	
	K.1	NEW	K.1	NEW
Bitmap size	2550×3300			
Compressed size (bytes)	338,776	277,617	707,819	258,765
Bits per pixel	0.32	0.26	0.67	0.25
Compression ratio	1 : 75	1 : 91	1 : 36	1 : 98

The Fuzzy Text Problem

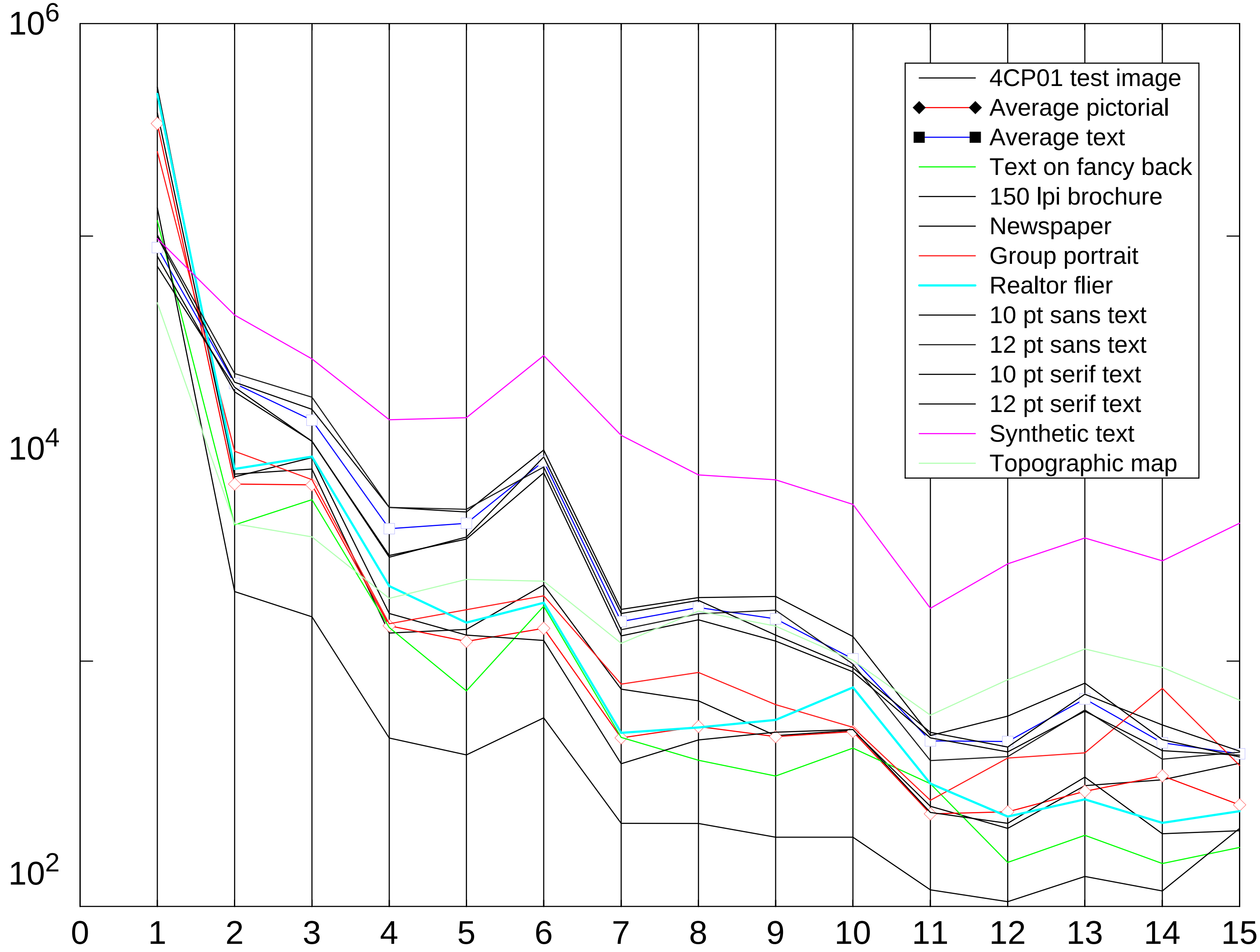
Cost reduction introduces problems such as

- sensor misalignment
- optical blur and electric cross-talk
- halftoning at low resolution

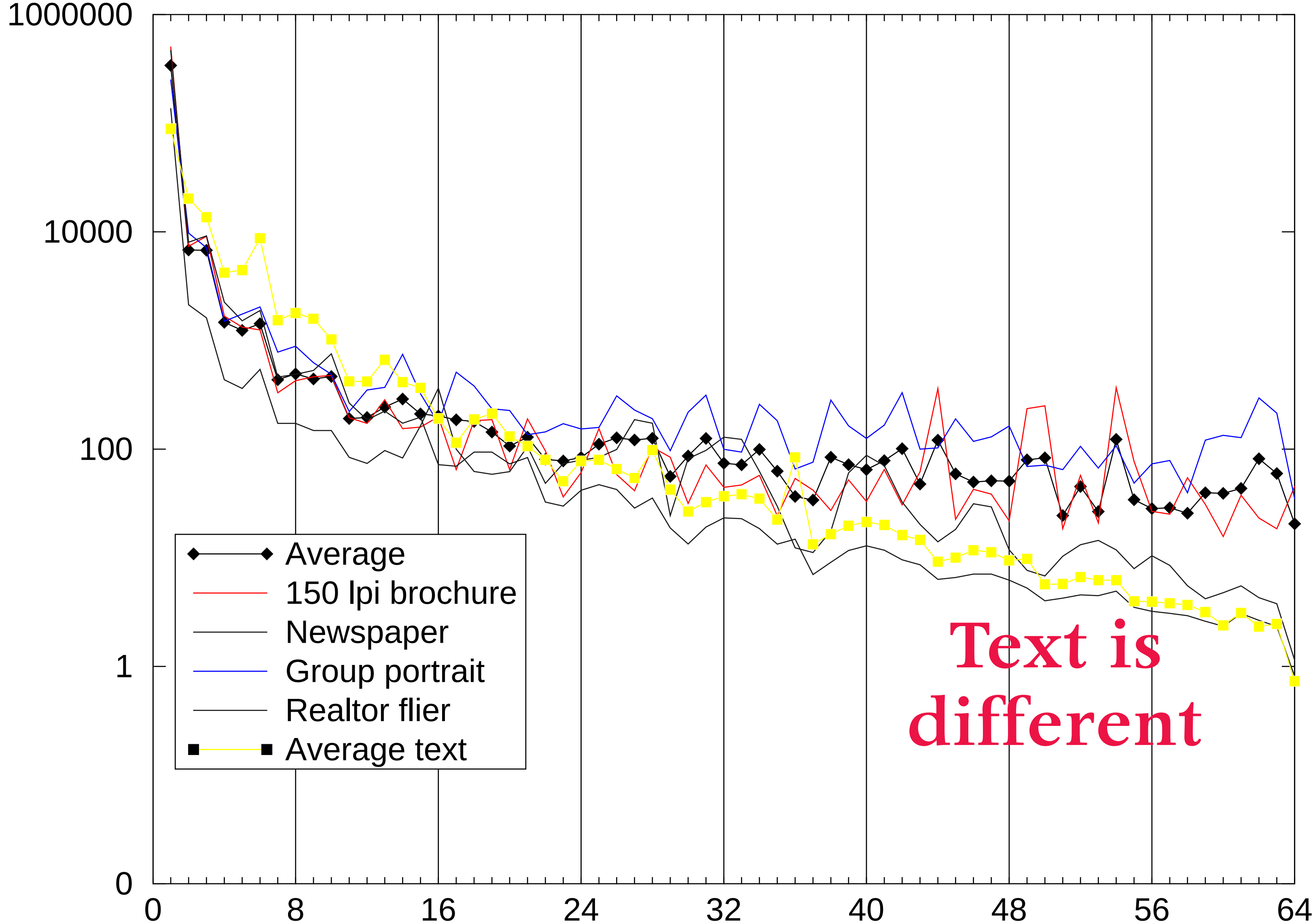
Comparison of L* Energy in all Images



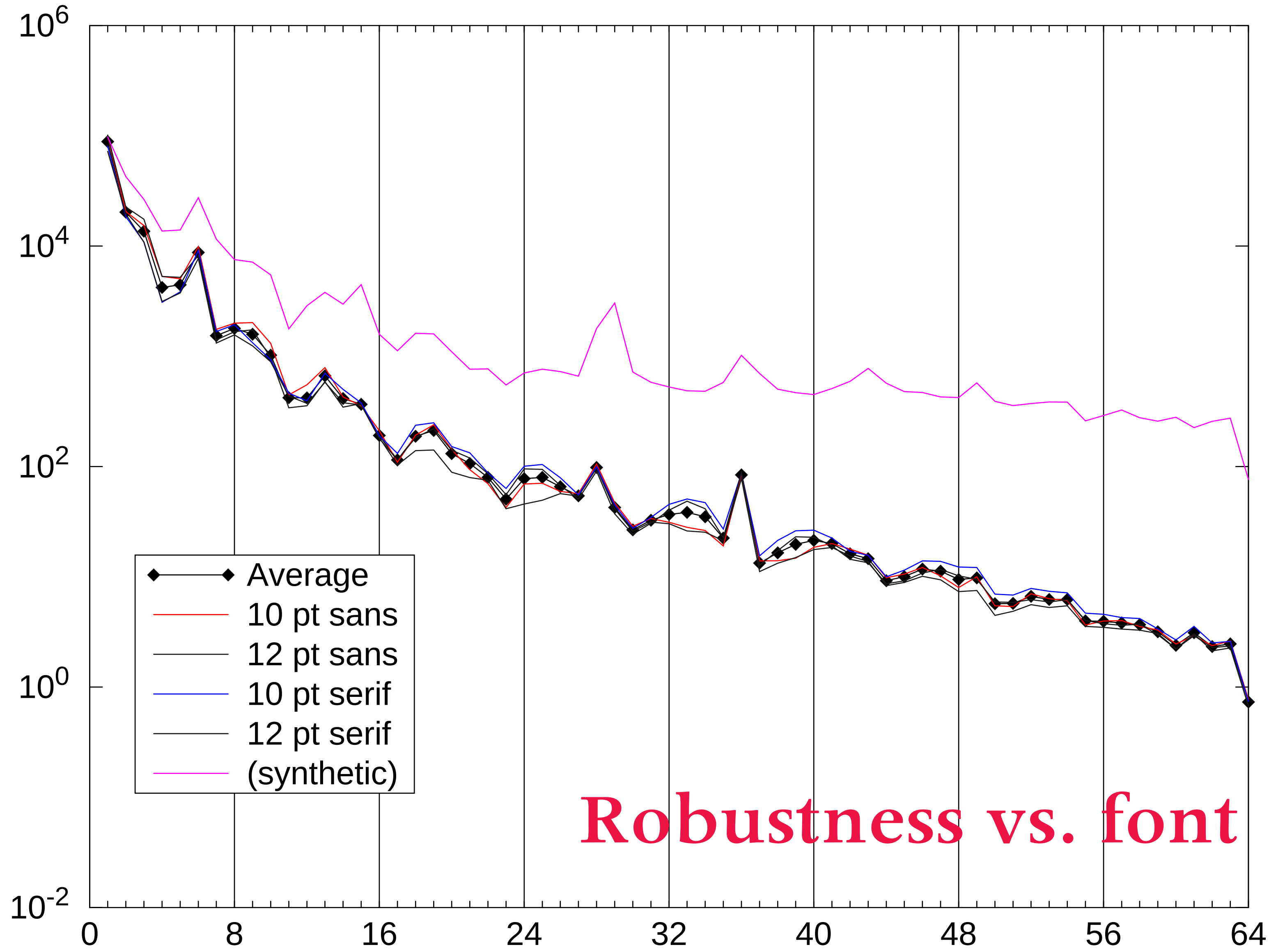
Comparison of Low Frequency Energy



L* Energy in Pictorial Images

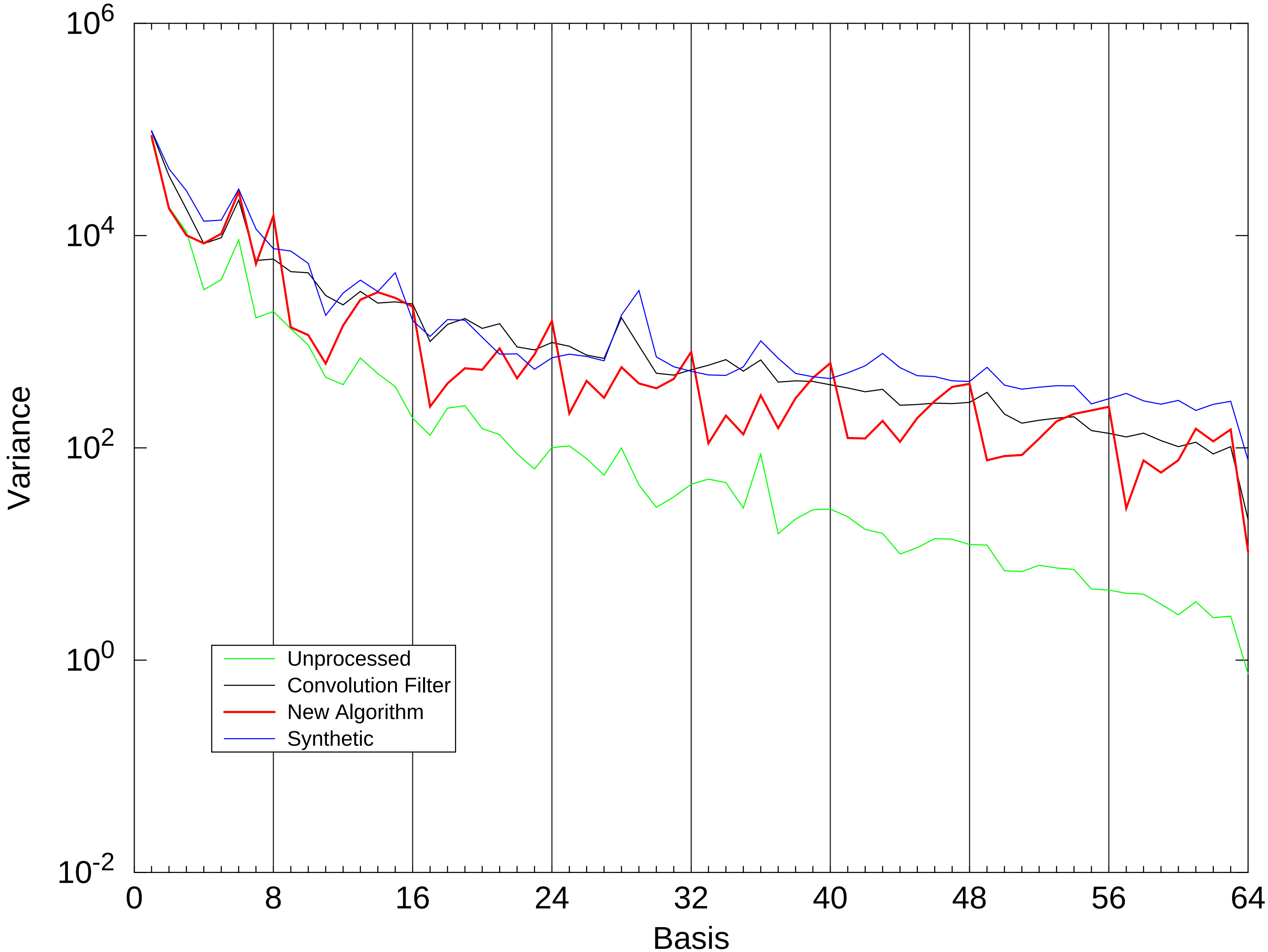


L* Energy in Text Images

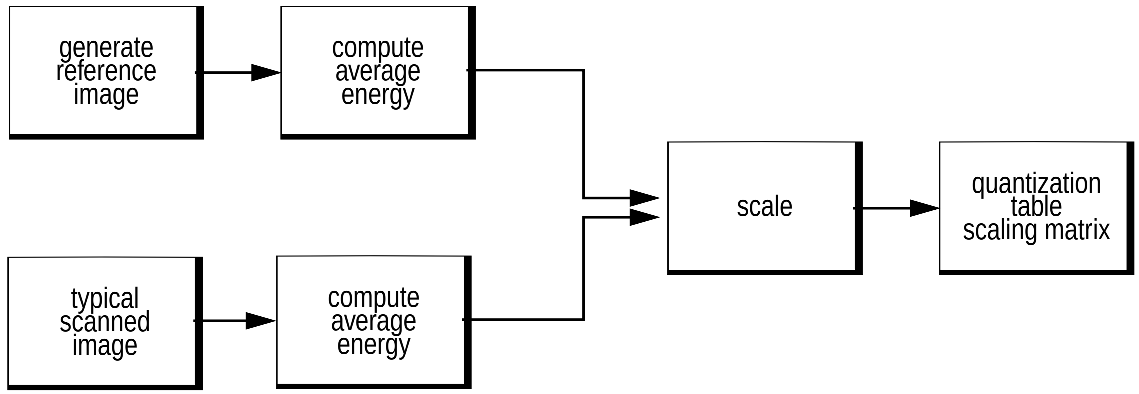


Robustness vs. font

Enhancement of 10 Point Serif Text



Sharpening an Image in the JPEG Domain During the Encoding



- Edge sharpening is achieved by using the *original* DQT for the encoding, while at the receiving fax machine the decoder uses the *scaled* DQT matrix
- The *scaled* matrix is the one included in the JPEG file
- Since only the *scaled* DQT is transmitted the original matrix and the actual factors can remain a trade secret

Huffman Tables

- Example Huffman Table is for perceptually lossless compression
- Color facsimile based on CIELAB color space

- Start with test chart
- Ad hoc technique: all symbol probabilities less than 2^{-16} are set to 2^{-16}
- Improvement: 8% to 14%
- Only 0.5% compression rate loss for other images
- Average improvement: 11%

Conclusions

- Digital images with text are not very robust with respect to quantization errors
- Sufficient information should be preserved in documents to preserve image quality when they are printed
- If the bandwidth is used judiciously, documents can be compressed to a higher degree, allowing the use of better resolutions or shorter transmission times
  1. Encode color in a perceptually uniform color space such as CIELAB
  2. Compress the spatial information using JPEG
  3. Design custom DQT tables for your document type
  4. Design custom Huffman tables