Image Compression Based on Visual Saliency at Individual Scales

Stella X. Yu[1] Dimitri A. Lisin[2]

[1]Computer Science Department
Boston College
Chestnut Hill, MA

[2]VideoIQ, Inc.
Bedford, MA

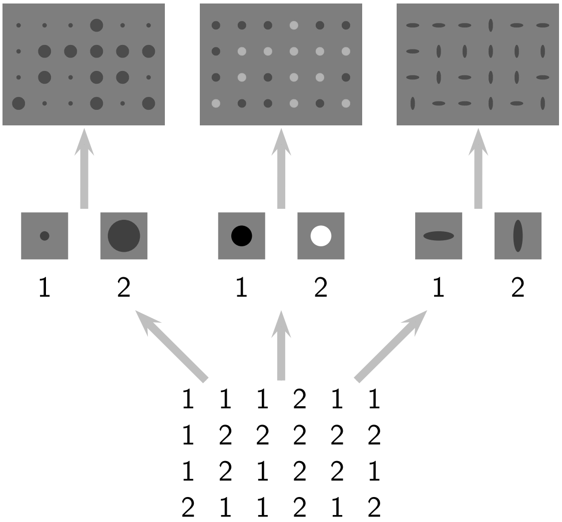
November 30, 2009

Motivation

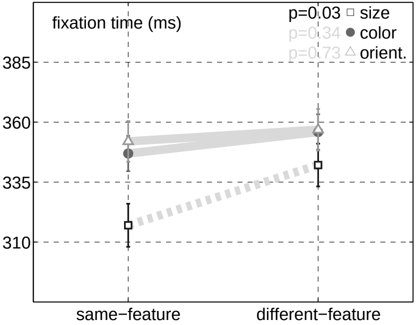
- Standard algorithms for lossy image compression:
  - reduce entropy while minimizing difference from the original
  - evaluate perceptual quality afterward
- Problem:
  - similarity to the original does not guarantee visual quality
- Goal:
  - reduce entropy while preserving salient regions

Looking Without Seeing

Experiments with Synthetic Images

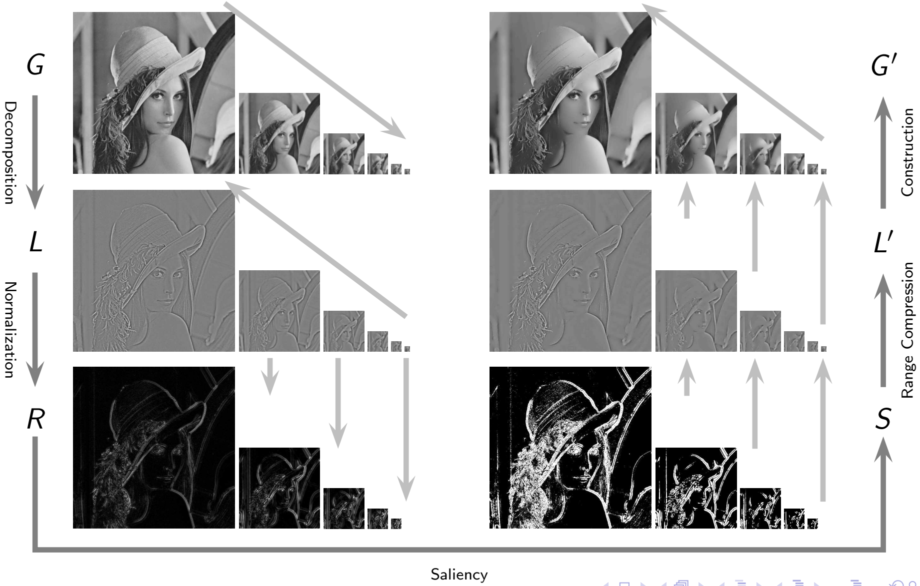


Results



- Preview benefits for size
- No preview benefits for intensity or orientation
- Conclusion: visual attention is scale-selective

Multi-scale Saliency-Based Compression



Adaptive Binning

- Value-adaptive binning
  - low values of the Laplacian are smoothed more than high values
- Scale-adaptive binning
  - range of the Laplacian naturally decreases with scale
  - finer scales are smoothed more than coarser scales

Test Images

Comparison with Signal-Based Compression

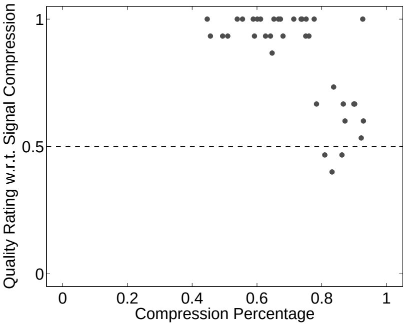
signal-based compression

our compression

Evaluation of Visual Quality

- 12 images, 3 compression levels, 15 (human) subjects
- Forced choice between a pair of images
- Brief exposure (1.2 s)
- Quality measured by percentage of votes

visual quality comparison:

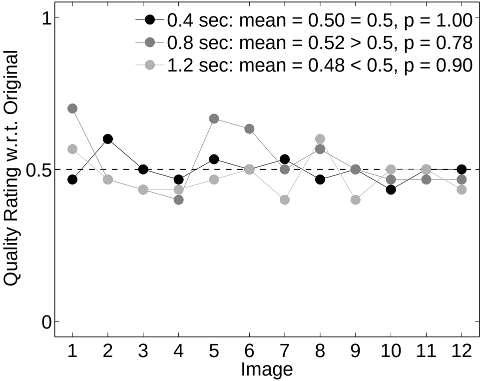


Comparison with Wavelet Compression

Daubechies wavelet compression

our compression

Comparison with the Originals



Compressed vs. Original

original image

our compression

Compressed vs. Original

original image

our compression

Summary

- Human vision study suggests scale-selectivity of attention
- Compression should preserve salient features at all scales
- Laplacian pyramid is used both as signal representation and as a saliency measure
- Range compression results in adaptive binning, reducing entropy while preserving visual quality
- Our algorithm can even enhance the visual quality of some images

Multi-scale Saliency Map

p specifies the fraction of pixels to be considered maximally salient
scaling factor α controls saliency sharpness
saliency map S at scale s is computed from L using a sigmoid with threshold m:

$$S_s = \left(1 + e^{-\frac{R_s - m_s}{\alpha}}\right)^{-1}, \; s = 1 : n \qquad (1)$$

$$m_s = \arg\left\{\sum_i S_s(i; m_s) = p \cdot \sum_i 1\right\}, R_s = \frac{|L_s|}{\max(|L_s|)} \qquad (2)$$

Entropy Reduction by Range Filtering

Given S, neighborhood radius r and range sensitivity factor β, generate a new Laplacian pyramid L' by spatially-variant range filtering of L:

$$L'_s(i) = \frac{\sum_{j \in N(i,r)} L_s(j) \cdot W_s(i,j)}{\sum_{j \in N(i,r)} W_s(i,j)}, L'_{n+1} = L_{n+1}, \; s = 1 : n \qquad (3)$$

$$W_s(i,j) = e^{-\frac{(L_s(i) - L_s(j))^2}{2\Theta_s(i)}}, \qquad (4)$$

$$\Theta_s(i) = (1 - S_s(i)) \cdot \left(\frac{\max(L_s) - \min(L_s)}{\beta} \right)^2 \qquad (5)$$