September 26, 2016

Visual Attention Guided Quality Assessment for Tone Mapped Images Using Scene Statistics

Debarati Kundu and Brian L. Evans

The University of Texas at Austin

The University of Texas at Austin
Electrical and Computer Engineering
Cockrell School of Engineering

WNCG Wireless Networking & Communications Group

Embedded Signal Processing Laboratory
ESPL
The University of Texas at Austin

Introduction

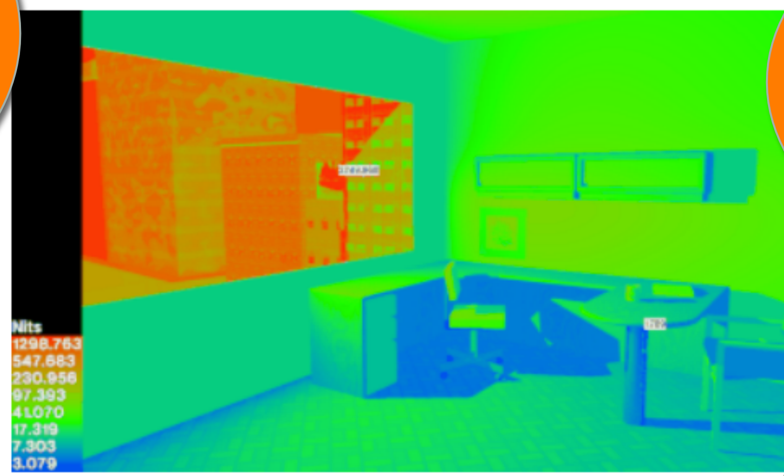
- Scene luminance varying from 10^{-4} to 10^{6} cd/m^{2} [Narwaria2013]
- High dynamic range (HDR) images preserve more detail

  - HDR picture capture (e.g. smart phones and DSLR cameras)
  - HDR video displays for home (e.g. Samsung)
  - HDR streaming content (e.g. Amazon Video and Netflix)
  - HDR graphics rendering (e.g. Unreal and CryEngine)

Tonemapping Operators [Larson1997]

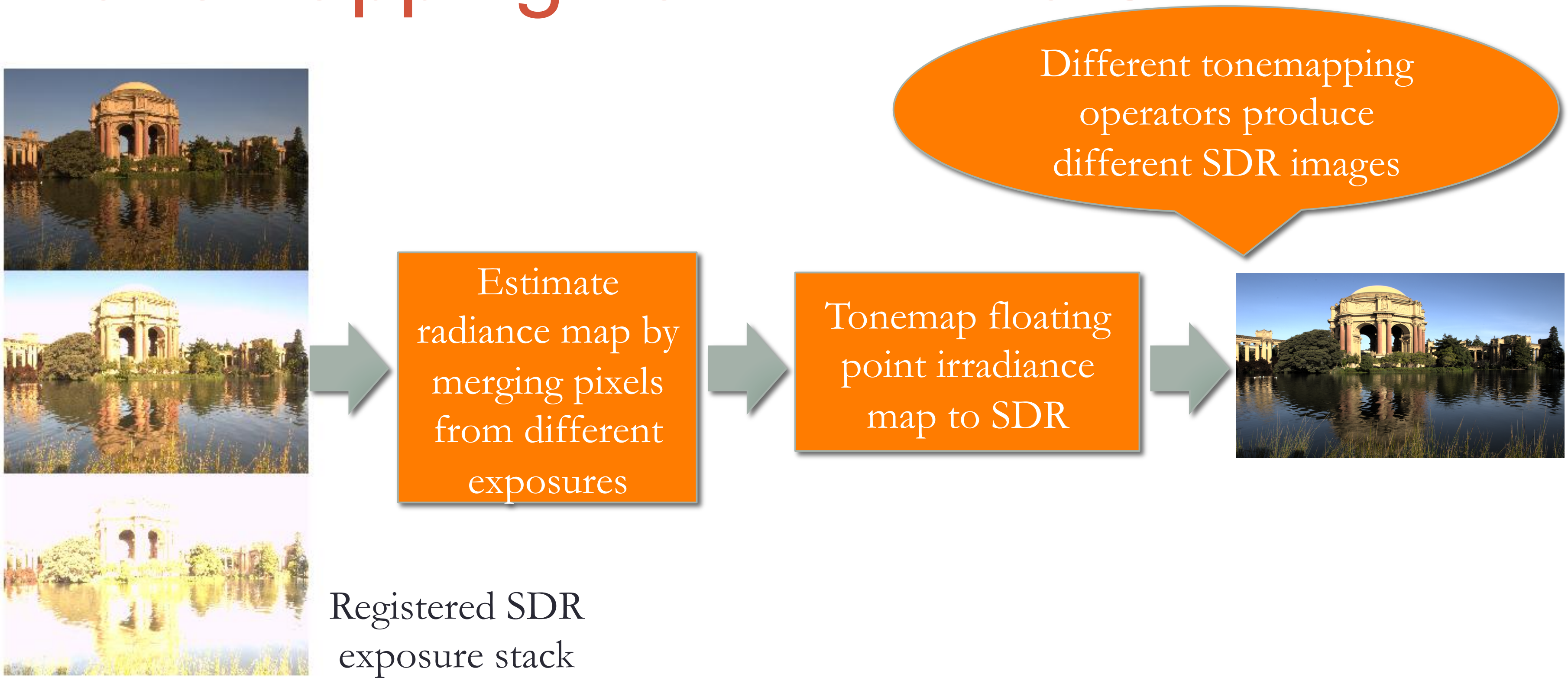
Uniformly spaced quantization of luminance overexposes the view through the window

World luminance values for a window office in candelas per meter squared



Luminance mapped to preserve visibility of both indoor & outdoor features using non-linear tonemapping

Tonemapping from HDR to SDR



- Propose three image quality assessment (IQA) algorithms
- Evaluate HDR radiance map and tonemapped SDR image

Tone Mapped Quality Index [Yeganeh2013]

- Overall Tone Mapped Quality Index
  - $a = 0.8012$, $\gamma = 0.3046$ and $\delta = 0.7088$

$$Q = aS^{\gamma} + (1-a)N^{\delta}$$

- Structural fidelity (*S*) of HDR and tonemapped SDR image
  - Structural similarity with penalty for large change in signal strength
  - *Pooling:* Average modified SSIM on 11 x 11 windows
  - Combine structural fidelity at each of five scales
- Naturalness (*N*) of tonemapped SDR image
  - Compute its global mean *m* and global standard deviation *d*
  - P_m and P_d are fits for global means & standard deviations for 3000 SDR natural images

$$N = \frac{P_m(m)\, P_d(d)}{\max\{P_m(m),\, P_d(d)\}} = \min\{P_m(m),\, P_d(d)\}$$

More Info

Change #1: Naturalness Measure

- Natural scene statistics (NSS) approach for IQA
  - Statistics of pristine images occur irrespective of content
  - Statistics of images with distortions deviate from scene statistics
- Mean subtracted contrast normalized pixels for image $I(i,j)$

[Ruderman1993]

$$\hat{I}(i,j) = \frac{I(i,j) - \mu(i,j)}{\sigma(i,j) + 1}$$

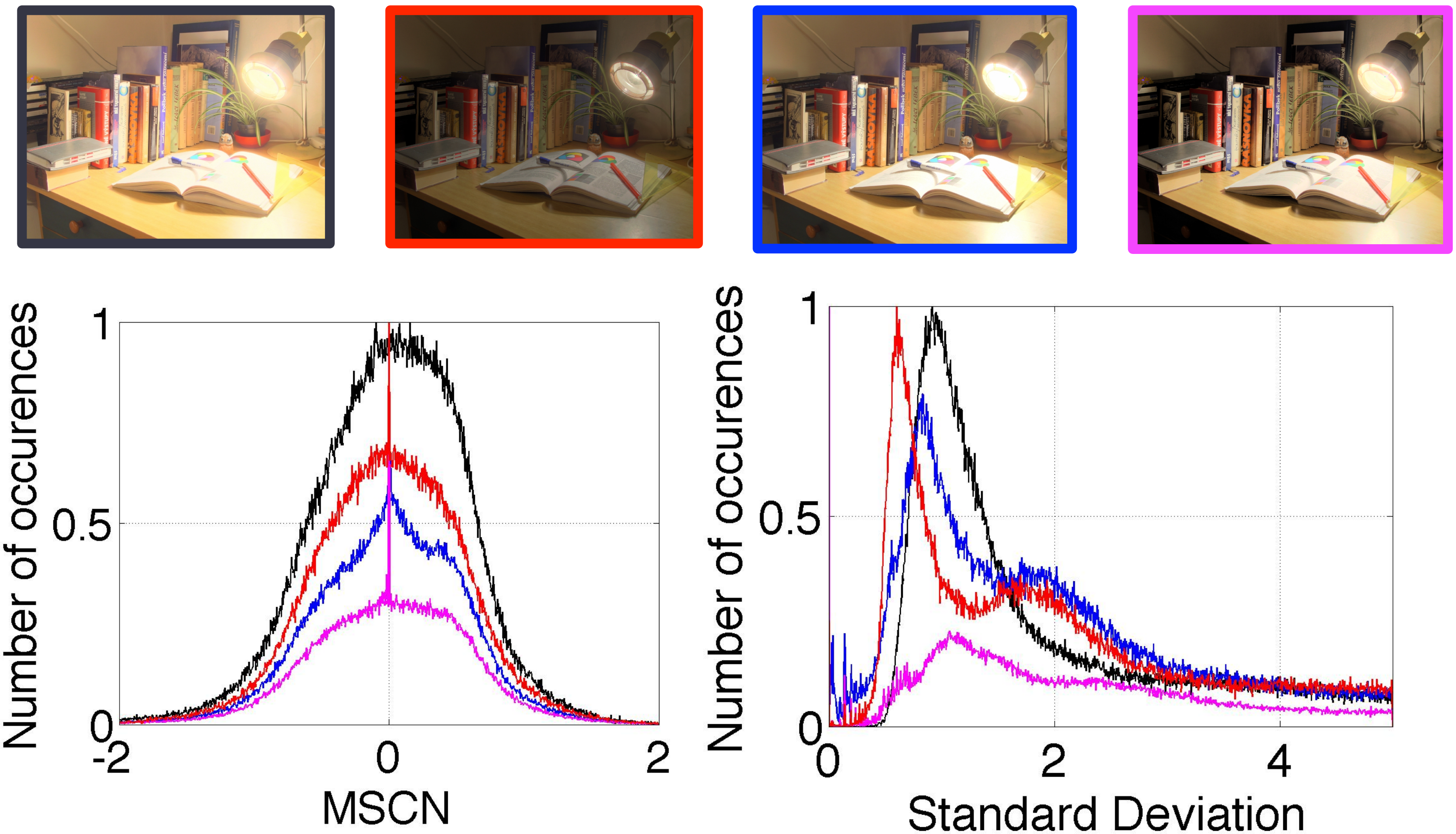
At pixel (i, j), use 11x11 window and uniform Gaussian filter (σ = 1.17)

$$\mu(i,j) = \sum_{k=-K}^{K} \sum_{l=-L}^{L} w_{k.l} I(i+k, j+l)$$ is weighted mean

$$\sigma(i,j) = \sqrt{\sum_{k=-K}^{K} \sum_{l=-L}^{L} w_{k.l} \left[I(i+k, j+l) - \mu(i,j) \right]^2}$$ is weighted standard deviation

MSCN models divisive normalization in retina

Tonemappings of Same Scene



MSCN coefficient distribution and σ-field distribution for different tonemapping operators

Proposed Naturalness Measure

- TMQI combines structural fidelity (S) and naturalness (N)

$$Q = aS^{\gamma} + (1-a)N^{\delta}$$

- Proposed naturalness measure based on scene statistics

$$Q = aS^{\gamma} + \frac{1}{2}(1-a)\beta^{\delta_1} + \frac{1}{2}(1-a)\phi^{\delta_2}$$

β: Exponent of generalized Gaussian fit of MSCN pixels of tonemapped SDR image

ϕ: Standard deviation of σ-field of tonemapped SDR image

a = 0.8012, γ = 0.3046 and $\delta_1 = \delta_2 = \delta$ = 0.7088 (same as in TMQI)

- Used in all three proposed IQA algorithms

Change #2: Pooling Approach

- Average pooling gives same importance to every pixel
- Information Maximization TMQI [Nasrinpour2015]
- Propose non-uniform pooling strategies using scene statistics
  - σ-map gives measures of edge magnitude and high contrast regions
  - Local entropy indicates local randomness (contrast)
  - Itti and Koch's saliency approach generalized for HDR images [Petit2009]

Tonemapped image

Structural fidelity map

Structural fidelity with pooling

TMQI Database [Yeganeh2013]

- 15 HDR source images, each mapped to SDR w/ 8 tonemaps
  - Subjects ranked 8 SDR images for every HDR source image
  - Correlated predicted and subjective ranks of tonemapped images
  - Median of correlation computations shown below

Full Reference IQA Algorithm	SROCC	PLCC	KCC	Time (s)
Proposed TMQI-NSS-σ pooling	**0.8810**	**0.9439**	**0.7857**	0.32
Proposed TMQI-NSS-Entropy pooling	**0.8810**	0.9438	0.7143	1.28
Proposed SHDR-TMQI pooling from [Petit2009]	**0.8810**	0.9346	0.7143	0.80
FSITM-TMQI [Nafchi2014]	0.8571	0.9230	**0.7857**	0.94
STMQI [Nasrinpour2015]	0.8503	0.9382	0.7638	1.54
TMQI-II [Ma2015]	0.8333	0.8790	0.7143	**0.20**
Feature Similarity Index for Tone-Mapped Images (FSITM) [Nafchi2014]	0.8333	0.8948	0.7143	0.47
TMQI [Yeganeh2013]	0.8095	0.9082	0.6429	0.52

HDR-JPEG Database [Narwaria2013]

- 10 source HDR images, each has 14 degraded versions
  - JPEG encoding at 7 different bit rates
  - SSIM and MSE used to design HDR->SDR and SDR->HDR mappings
  - 27 subjects rated individual HDR images on HDR displays on 1-5 scale

Full Reference IQA Algorithm	SROCC	PLCC	KCC	Time (s)
Proposed SHDR-TMQI pooling from [Petit2009]	**0.8510**	0.8533	0.6700	3.00
Proposed TMQI-NSS-σ pooling	0.8485	0.8520	0.6659	1.65
Proposed TMQI-NSS-Entropy pooling	0.8454	**0.8645**	**0.6719**	6.74
TMQI [Yeganeh2013]	0.7947	0.8057	0.6127	3.45
FSITM-TMQI [Nafchi2014]	0.6300	0.6584	0.4762	8.35
TMQI-II [Ma2015]	0.5096	0.5137	0.3642	**1.34**
Feature Similarity Index for Tone-Mapped Images (FSITM) [Nafchi2014]	0.4720	0.5167	0.3422	5.26
STMQI [Nasrinpour2015]	0.3464	0.3244	0.2449	12.00

Conclusion

- Perceptually-guided pooling boosts correlation with human subjective ratings vs. average pooling
- Pooling using σ-map has good correlation vs. runtime tradeoff
- Software: http://signal.ece.utexas.edu/~bevans/HDRImaging/

More Recent Work

- ESPL-LIVE HDR Image Database of 1800+ HDR pictures

http://signal.ece.utexas.edu/~debarati/ESPL_LIVE_HDR_Database

  - Crowdsourced study with 5000 observers and 300,000 opinion scores
  - Proposed and evaluated no-reference IQA algorithms for HDR images
- Joint effort with D. Ghadiyaram and A. C. Bovik, UT Austin

References

[Larson1997] G. W. Larson, H. Rushmeier, and C. Piatko. “A Visibility Matching Tone Reproduction Operator for High Dynamic Range Scenes”, *IEEE Trans. on Visualization and Computer Graphics* 3, 4, Oct. 1997, pp. 291-306.

[Ma2015] Kede Ma; Yeganeh, H.; Kai Zeng; Zhou Wang, "High Dynamic Range Image Compression by Optimizing Tone Mapped Image Quality Index,” *IEEE Trans. on Image Processing*, vol.24, no.10, pp.3086-3097, Oct. 2015

[Nafchi2014] H. Ziaei Nafchi, A. Shahkolaei, R. Farrahi Moghaddam, and M. Cheriet, “Fsitm: A feature similarity index for tone-mapped images,” *IEEE Signal Processing Letters*, vol. 22, no. 8, pp. 1026-1029, Aug. 2015.

[Narwaria2013] M. Narwaria, M. Perreira Da Silva, P. Le Callet, and R. Pepion, “Tone mapping-based high-dynamic-range image compression: study of optimization criterion and perceptual quality,” *Optical Engineering*, vol. 52, no. 10, Oct 2013.

[Nasrinpour2015] 1 H. R. Nasrinpour and N. D. Bruce, “Saliency weighted quality assessment of tone-mapped images,” *Proc. IEEE Int. Conf. Image Proc.*, Sep. 2015.

[Petit2009] J. Petit, R. Brémond, and J.-P. Tarel, “Saliency maps of high dynamic range images,” *Proc. ACM Symp. Appl. Perception in Graphics & Visualization*, 2009.

[Ruderman1993] D. L. Ruderman and W. Bialek, “Statistics of natural images: Scaling in the woods,” *Proc. Neural Info. Processing Sys. Conf. and Workshops*, 1993.

[Yeganeh2013] H. Yeganeh and Z. Wang, “Objective quality assessment of tone-mapped images,” *IEEE Trans. on Image Processing* , vol. 22, no. 2, pp. 657-667, Feb 2013.

Questions?


THANK
YOU
GRACIAS
ARIGATO
SHUKURIA
DANKSCHEEN
BİYAN
SHUKRIA
TINGKI
YAQHANYELAY
TASHAKKUR ATU
SUKSAMA
EKHMET
GRAZIE
MEHRBANI
MAAKE
KOMAPSUMNIDA
PALDIES
GOZAIMASHITA
EFCHARISTO
JUSPAXAR
BOLZÏN
MERCI

Multi-Exposure Fusion

SDR

Merge exposure stack directly to get fused image

HDR but in SDR format

$$Y(i) = \sum_{k=1}^{K} W_k(i) X_k(i)$$

*i*th pixel index

*k*th exposure image

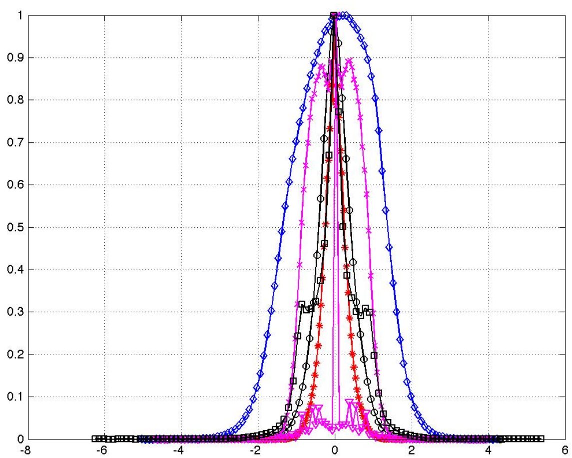
$X_k(i)$ luminance

$W_k(i)$ weight for perceptual importance of exposure level *k*

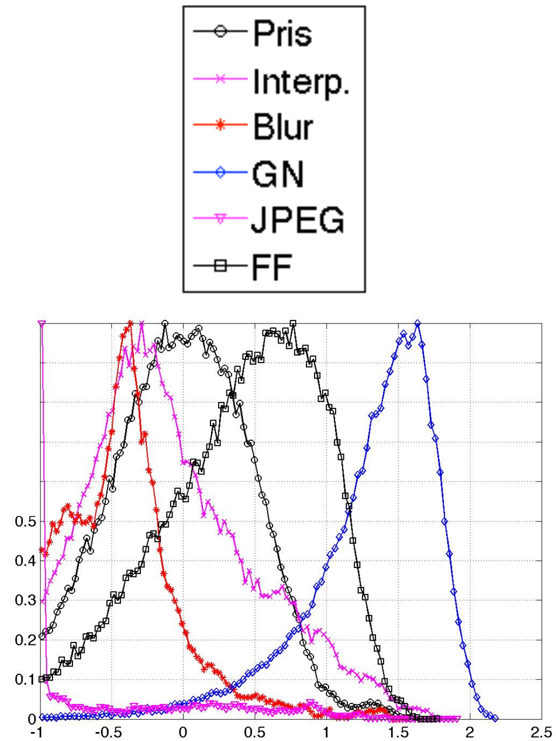
- Registered exposure stack of *K* images
  - Standard dynamic range (SDR) images
  - Requires camera calibration and motion compensation

Distorted Image Statistics

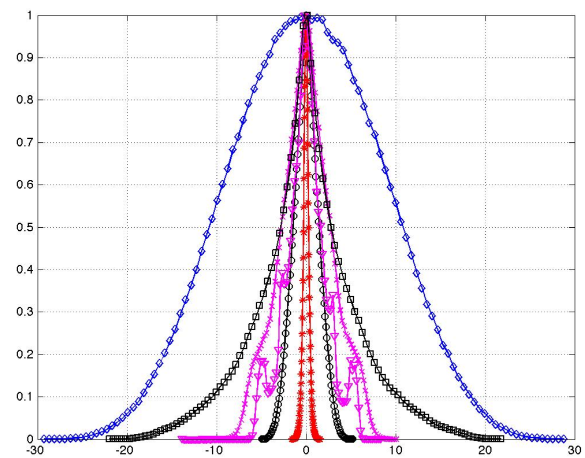
- Different distortions affect scene statistics characteristically
- Used for distortion classification and blind quality prediction



MSCN Coefficients



Curvelet Coefficients



Steerable Pyramid Wavelet Coefficients

Back

Tone Mapped Quality Index [Yeganeh2013]

- Tonemapping meant to change local intensity & contrast
- Structural fidelity modifies Structural Similarity (SSIM)
  - Penalizes large change in strength in HDR vs. SDR image patch
  - Local standard deviations nonlinearly mapped via Gaussian CDF
    - Significant signal strength mapped to 1
    - Insignificant signal strength mapped to 0

$$p(s) = \frac{1}{\sqrt{2\pi}\theta_s} \int_{-\infty}^{s} \exp\left[-\frac{(x-\tau_s)^2}{2\theta_s^2}\right] dx$$

- Structural fidelity computation over five scales
- Naturalness measure of tonemapped SDR image
  - Distribution of global means in 3000 natural images

$$P_m(m) = \frac{1}{\sqrt{2\pi}\sigma_m} \exp\left[-\frac{m-\mu_m}{2\sigma_m^2}\right]$$

  - Distribution of global standard deviations in 3000 natural images

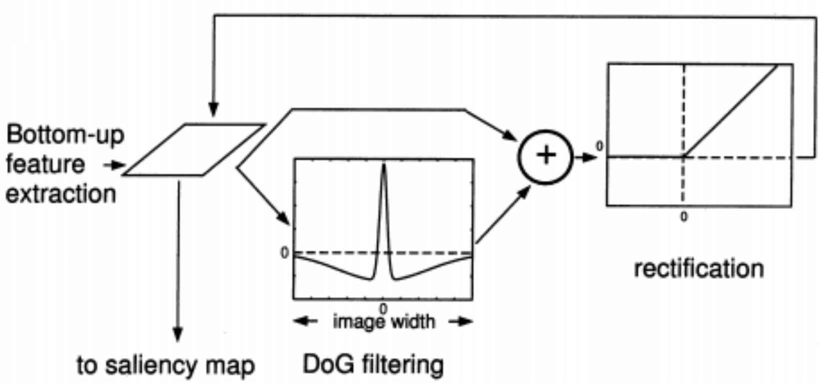
$$P_d(d) = \frac{(1-d)^{\beta_d - 1} d^{\alpha_d - 1}}{\mathrm{B}(\alpha_d, \beta_d)}$$

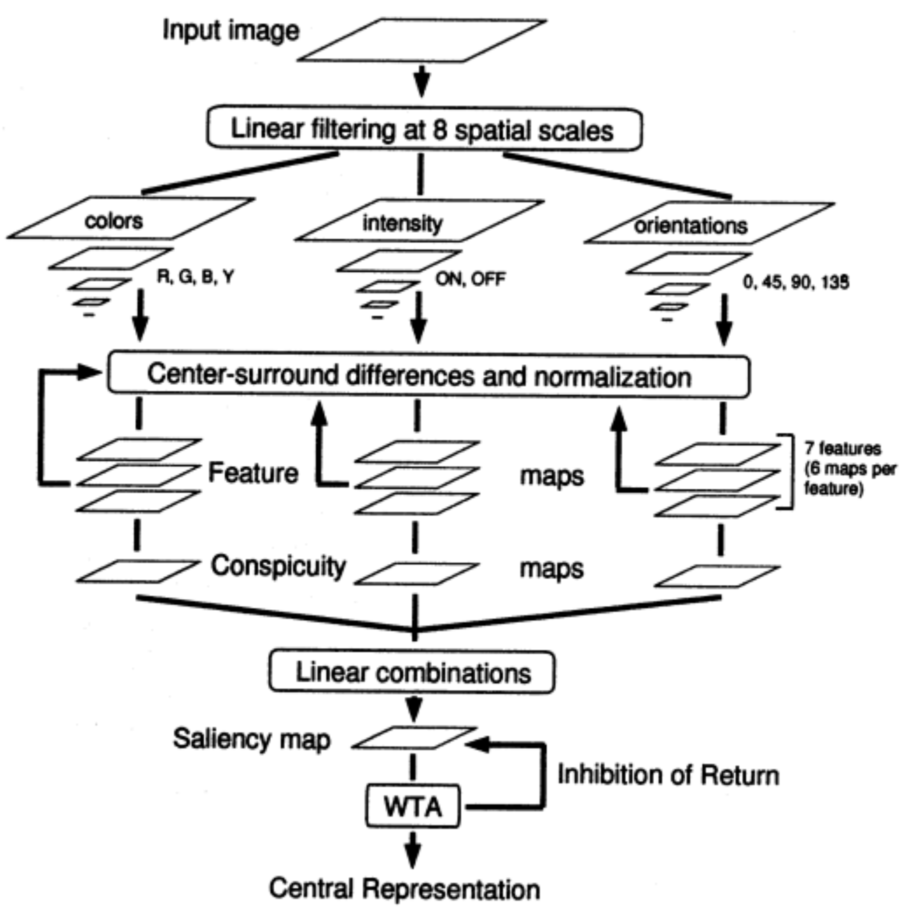
Back

Itti and Koch's Saliency

Back

- Different scales
  - Implemented as Gaussian Pyramid
- Center Surround mechanism
  - Implemented with DoG
- LPF repeated over multiple scales
- 3 scales, 4 orientations used





Generalized Gaussian Density

- **GGD** $p_g(r) = \frac{\beta}{2\sigma\Gamma(\beta^{-1})} \exp(|r|/\sigma)^{\beta} \quad r \in \Re,\ \sigma,\beta > 0$

  includes the **special cases**

  $\beta = 1$ (**Laplacian** density)

  $\beta = 2$ (**Gaussian** density)

  $\beta = \infty$ (**uniform** density)

- Many authors observe **GGD behavior** of **bandpass image** signals
  - Wavelet coefficients
  - DCT coefficients
  - **Usually** reported that β » 1 but **varies** $(0.8 < \beta < 1.4)$

[A. C. Bovik, *EE381V Digital Video*, UT Austin, Spring 2015]

Calculating Correlations

- Spearman's Rank-Order Correlation Coefficient (SRCC)

$$SRCC = 1 - \frac{6\sum_{i=1}^{N} d_i^2}{N(N^2-1)}$$

  d_i is difference between ith image's ranks is subjective and objective evaluations

  N is number of rankings

- Kendall's correlation coefficient (KCC)

$$KCC = \frac{N_c - N_d}{0.5N(N-1)}$$

  N_c and N_d are the number of concordant (of consistent rank order) and discordant (of inconsistent rank order) pairs in the data set respectively

  N is number of rankings

- Pearson's Linear Correlation Coefficient (PLCC)

$$r = \frac{n\left(\sum_{i=1}^{n} x_i y_i\right) - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{\sqrt{\left(n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2\right)\left(n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y\right) y_i^2\right)}}$$

Back