NEAR-THRESHOLD PERCEPTUAL DISTORTION PREDICTION BASED ON OPTIMAL STRUCTURE CLASSIFICATION

Yucheng Liu and Jan P. Allebach
School of ECE, Purdue University
09/26/2016

Presenter:
Jan P. Allebach
allebach@purdue.edu

PURDUE UNIVERSITY

Outline

- Introduction to near-threshold local distortion in natural images
- Limitation of current methods
- Probabilistic model for near-threshold distortion
- Results
- Contributions

PURDUE UNIVERSITY

Outline

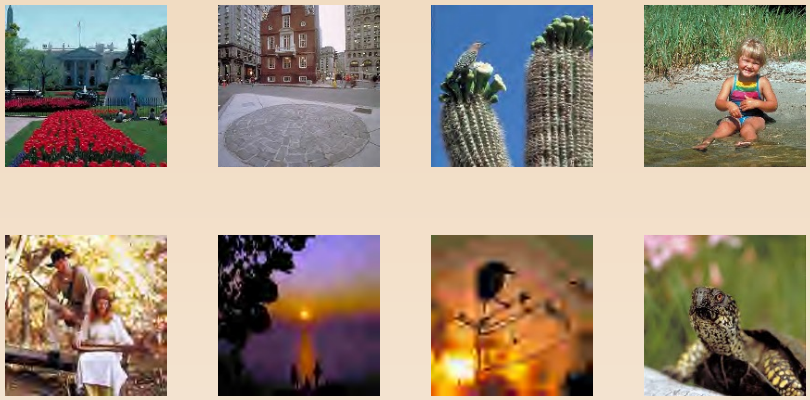
- Introduction to near-threshold local distortion in natural images
- Limitation of current methods
- Probabilistic model for near-threshold distortion
- Results
- Conclusion

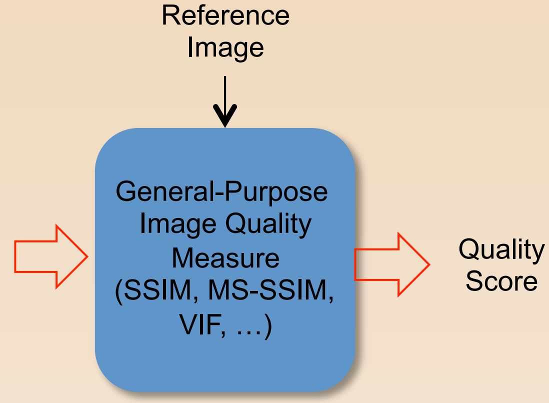
PURDUE UNIVERSITY

Near-Threshold Local Image Quality

- General-Purpose Image Quality Measure

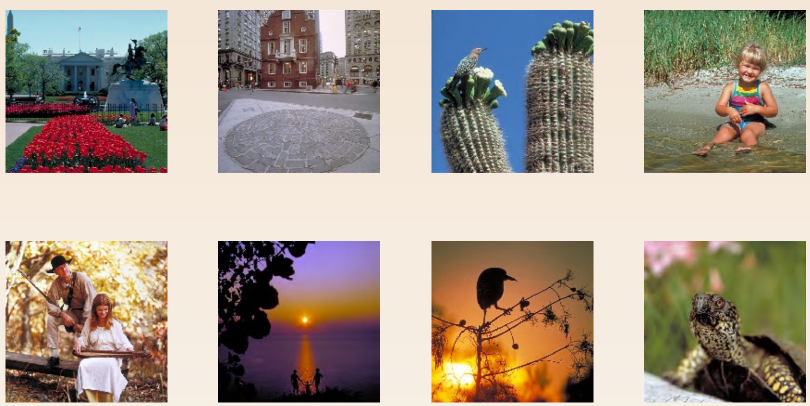
Distortion is present at both near-threshold and (mostly) supra-threshold level

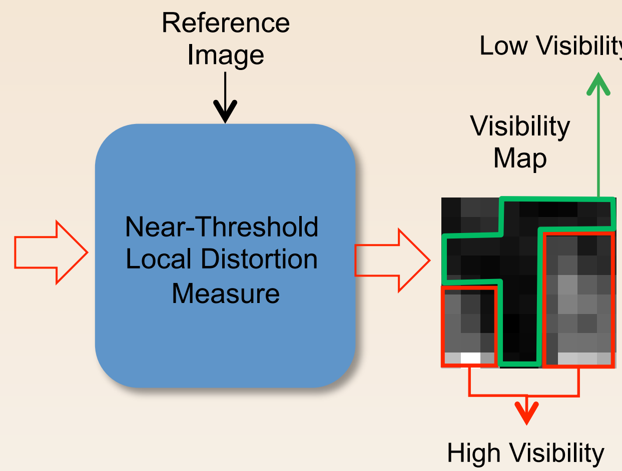




- Near-Threshold Local Image Quality Measure
  - » Potential applications: Image compression, digital watermarking

Distortion is present at below- and near-threshold level

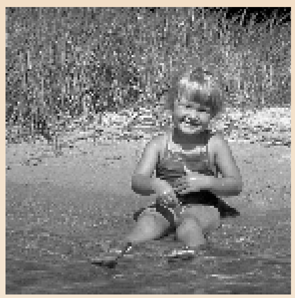




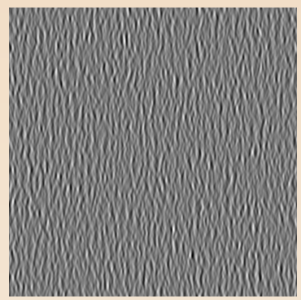
PURDUE UNIVERSITY

Problem Definition

- Goal: For a given reference image and distortion, predict the distortion visibility threshold (in RMS contrast) at each local region.

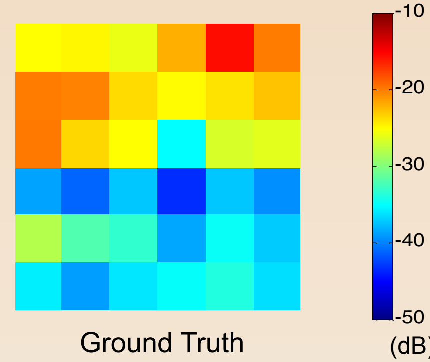


Reference Image



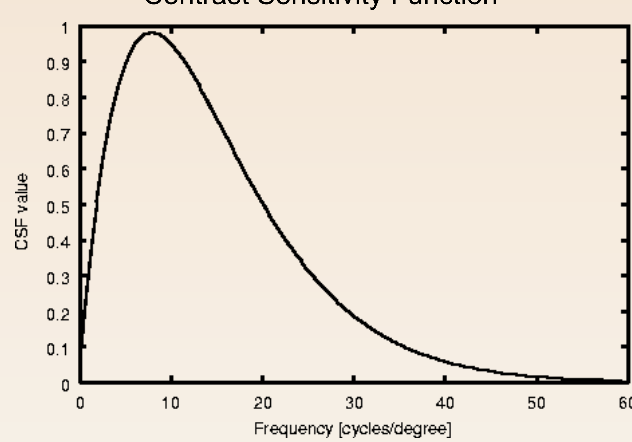
Error Image

Perceptual Model



Ground Truth Threshold Map

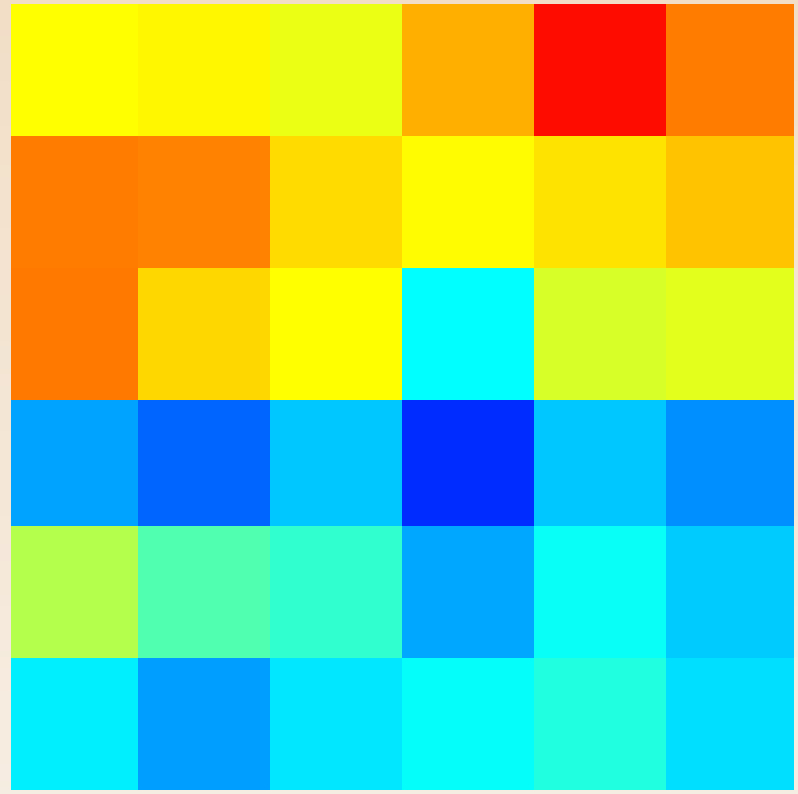
- Key properties of the HVS related to the near-threshold local distortion visibility:
  - » Multi-channel frequency response
  - » Contrast sensitivity
  - » Light Adaptation
  - » Contrast masking
  - » Structure familiarity



[J. L. Mannos and D. J. Sakrison'1974]

PURDUE UNIVERSITY

Detail From Previous Slide



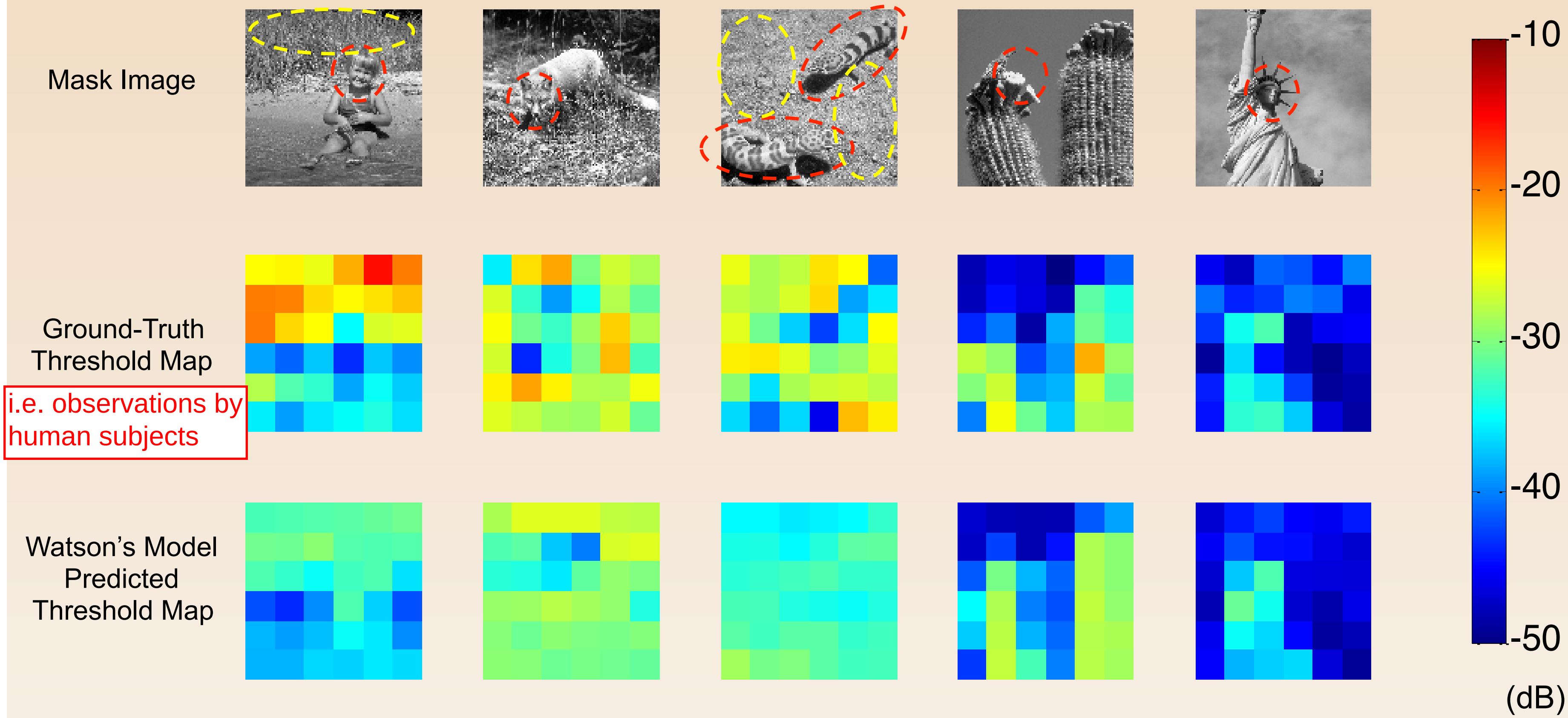
Outline

- Introduction to near-threshold local distortion in natural images
- Limitation of current methods
- Probabilistic model for near-threshold distortion
- Results
- Conclusion

PURDUE UNIVERSITY

Limitations of Gain-Control Based Models

- Gain-control based models are usually optimized on unnatural mask images, it cannot adapt flexibly to complicated natural image structures



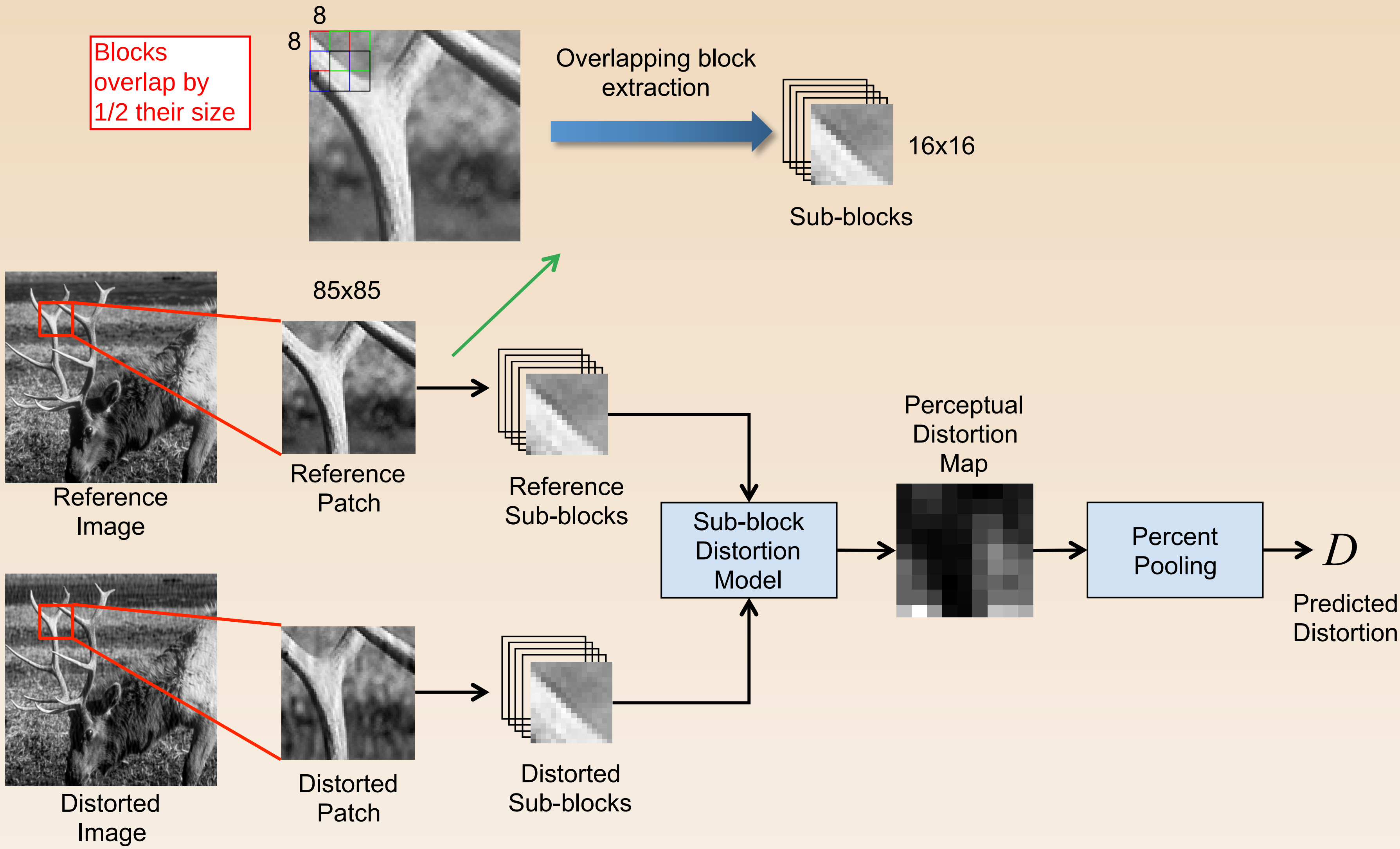
PURDUE UNIVERSITY

Outline

- Introduction to near-threshold local distortion in natural images
- Limitation of current methods
- Optimal classification based model for near-threshold distortion
- Results
- Conclusion

PURDUE UNIVERSITY

High Level Local Distortion Prediction Framework



Probabilistic Problem Formulation

- The perceived near-threshold distortion can be modeled as the probability that the local distortion is visible given the local distortion contrast, local lightness, and local structures

$$d = P(y=1 \mid \mathrm{L}^*, \mathbf{c}_{err}, \mathbf{s})$$

y = 1 means distortion is visible
y = 0 means distortion is not visible
y = 1/2 means distortion is near threshold

$$= \sum_{k=1}^{K} P(k_s = k \mid \mathrm{L}^*, \mathbf{c}_{err}, \mathbf{s}) \cdot P(y=1 \mid k_s = k, \mathrm{L}^*, \mathbf{c}_{err}, \mathbf{s})$$

Structure Class Label | Local Lightness | Distortion Contrast Feature | Local Structure Feature

- Model assumption:

$$P(k_s = k \mid \mathrm{L}^*, \mathbf{c}_{err}, \mathbf{s}) = P(k_s = k \mid \mathbf{s})$$

Given local structure, structure class is independent of local lightness and distortion contrast

$$P(y=1 \mid k_s = k, \mathrm{L}^*, \mathbf{c}_{err}, \mathbf{s}) = P(y=1 \mid k_s = k, \mathrm{L}^*, \mathbf{c}_{err})$$

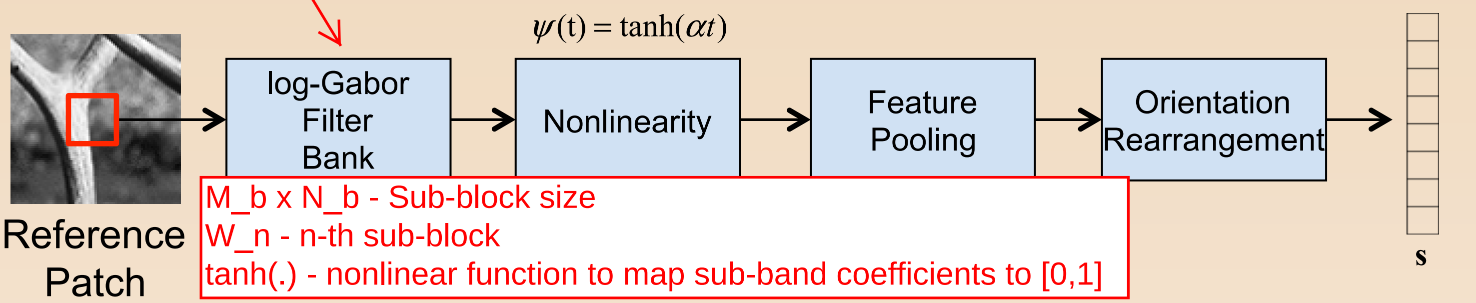
Given structure class, visibility of distortion is independent of local structure

Local Structure Feature Extraction

[22] D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," J. Opt. Soc. Am. A , pp. 2379–2397, 1987.

i - index of log-Gabor filter output
alpha - model parameter

$\psi(\mathrm{t}) = \tanh(\alpha t)$

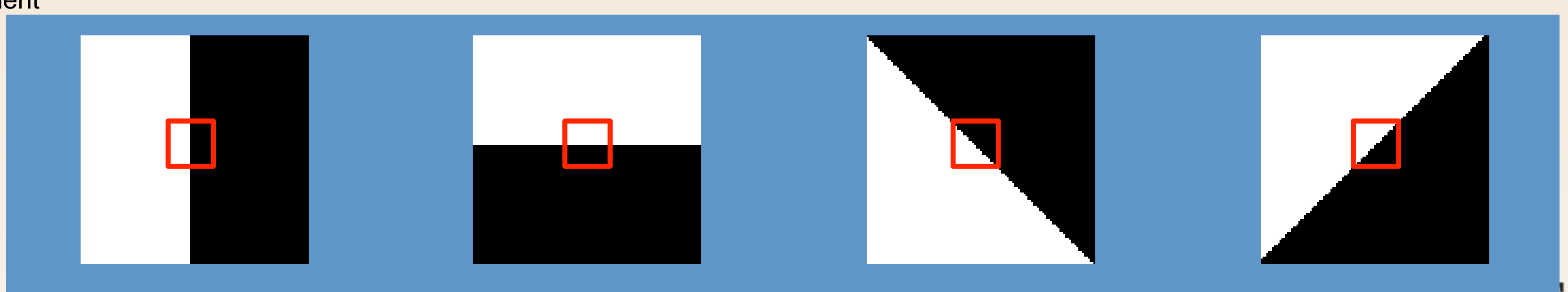


M_b x N_b - Sub-block size
W_n - n-th sub-block
tanh(.) - nonlinear function to map sub-band coefficients to [0,1]

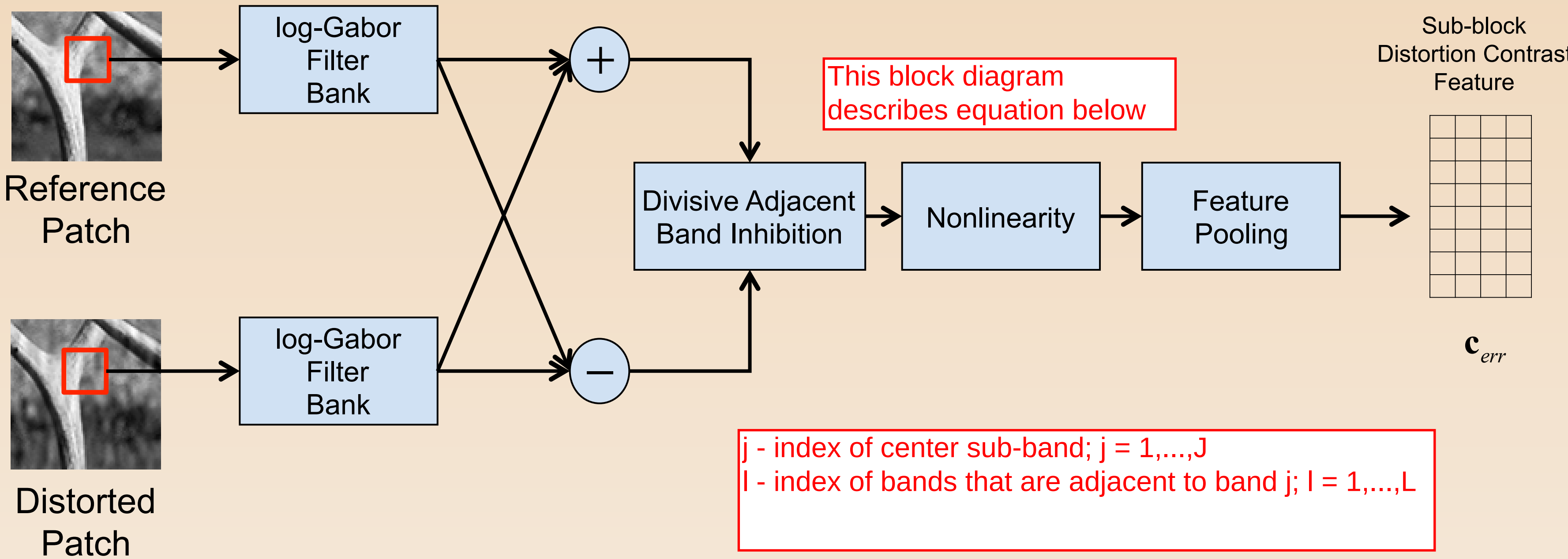
$$s_{m,n}(i) = \frac{1}{M_b \times N_b} \sum_{(x,y)\in W_n} |\tanh(z_m^r(i,x,y))|$$

- Structure feature is rotation invariant (similar to SIFT).
- Orientation rearrangement shifts the structure feature vector to prioritize the orientation component with strongest spectral energy magnitude at each radial frequency band.
- At prediction stage, spatial consistency is enforced for each sub-block structure feature.

W/ Orientation Rearrangement	$\mathbf{s}_1 = (0.64, 0.34, 0, 0.34)^T$	$\mathbf{s}_2 = (0.64, 0.34, 0, 0.34)^T$	$\mathbf{s}_3 = (0.64, 0.34, 0, 0.34)^T$	$\mathbf{s}_4 = (0.64, 0.34, 0, 0.34)^T$
W/O Orientation Rearrangement	$\mathbf{s}_1 = (0.64, 0.34, 0, 0.34)^T$	$\mathbf{s}_2 = (0, 0.34, 0.64, 0.34)^T$	$\mathbf{s}_3 = (0.34, 0.64, 0.34, 0)^T$	$\mathbf{s}_4 = (0.34, 0, 0.34, 0.64)^T$



Distortion Contrast Feature Extraction



- The distortion contrast feature depicts the masked error contrast in each frequency band

captures masking effect due to energy in adjacent bands

$$c_{m,n}(j,l) = \frac{1}{M_b \times N_b} \sum_{(x,y)\in W_n} \left| \frac{z_m^{L^{r*}}(j,x,y) - z_m^{L^{d*}}(j,x,y)}{z_m^{L^{r*}}(n_l^j,x,y) + z_m^{L^{d*}}(n_l^j,x,y)} \right|^{\rho}$$

models nonlinear contrast response of the HVS

- The set of adjacent bands $N^j = \{n_1^j, n_2^j, \ldots, n_L^j\}$ include subbands within 45 degrees, and 1.7 octaves of the center band.

PURDUE UNIVERSITY

Structure Classification Model

- The distribution of structure features is modeled by a Gaussian Mixture Model (GMM):

$$P(\mathbf{s}) = \sum_{k=1}^{K} P(\mathbf{s} \mid k_s = k) P(k_s = k)$$

$$= \sum_{k=1}^{K} \frac{\pi_k}{(2\pi)^{l/2} \mid \Sigma_k \mid^{1/2}} \exp\left(-\frac{1}{2} (\mathbf{s} - \boldsymbol{\mu}_k)^T \Sigma_k (\mathbf{s} - \boldsymbol{\mu}_k) \right)$$

- Parameters $\pi_k, \boldsymbol{\mu}_k, \text{ and } \Sigma_k$ can be obtained by solving the following maximum likelihood problem on the training data using EM algorithm:

$$(\pi^*, \boldsymbol{\mu}^*, \Sigma^*) = \underset{\pi, \boldsymbol{\mu}, \Sigma}{\operatorname{argmin}} \sum_{\mathbf{s} \in \Psi} -\log P(\mathbf{s})$$

- The classification probability can be computed as

$$P(k_s = j \mid \hat{\mathbf{s}}) = P_{j|\hat{\mathbf{s}}} = \frac{\dfrac{\pi_j^*}{\mid \Sigma_j^* \mid^{1/2}} \exp\left(-\frac{1}{2} (\hat{\mathbf{s}} - \boldsymbol{\mu}_j^*)^T \Sigma_j^* (\hat{\mathbf{s}} - \boldsymbol{\mu}_j^*) \right)}{\displaystyle\sum_{k=1}^{K} \frac{\pi_k^*}{\mid \Sigma_k^* \mid^{1/2}} \exp\left(-\frac{1}{2} (\hat{\mathbf{s}} - \boldsymbol{\mu}_k^*)^T \Sigma_k^* (\hat{\mathbf{s}} - \boldsymbol{\mu}_k^*) \right)}$$

Distortion Prediction Model

- The perceptual distortion prediction probability given the distortion contrast, luminance, and local structure type is defined as

model parameters

$$P(y=1 \mid k_s = k, L^*, \mathbf{c}_{err}) = \phi_k(L^*, \mathbf{c}_{err})$$
$$= S\left(b_k + \mathbf{w}_k^T\left(g\left(L^*\right)\mathbf{c}_{err}\right)\right)$$

where $S(t)=\dfrac{1}{1+e^{-t}}$ is the sigmoid function

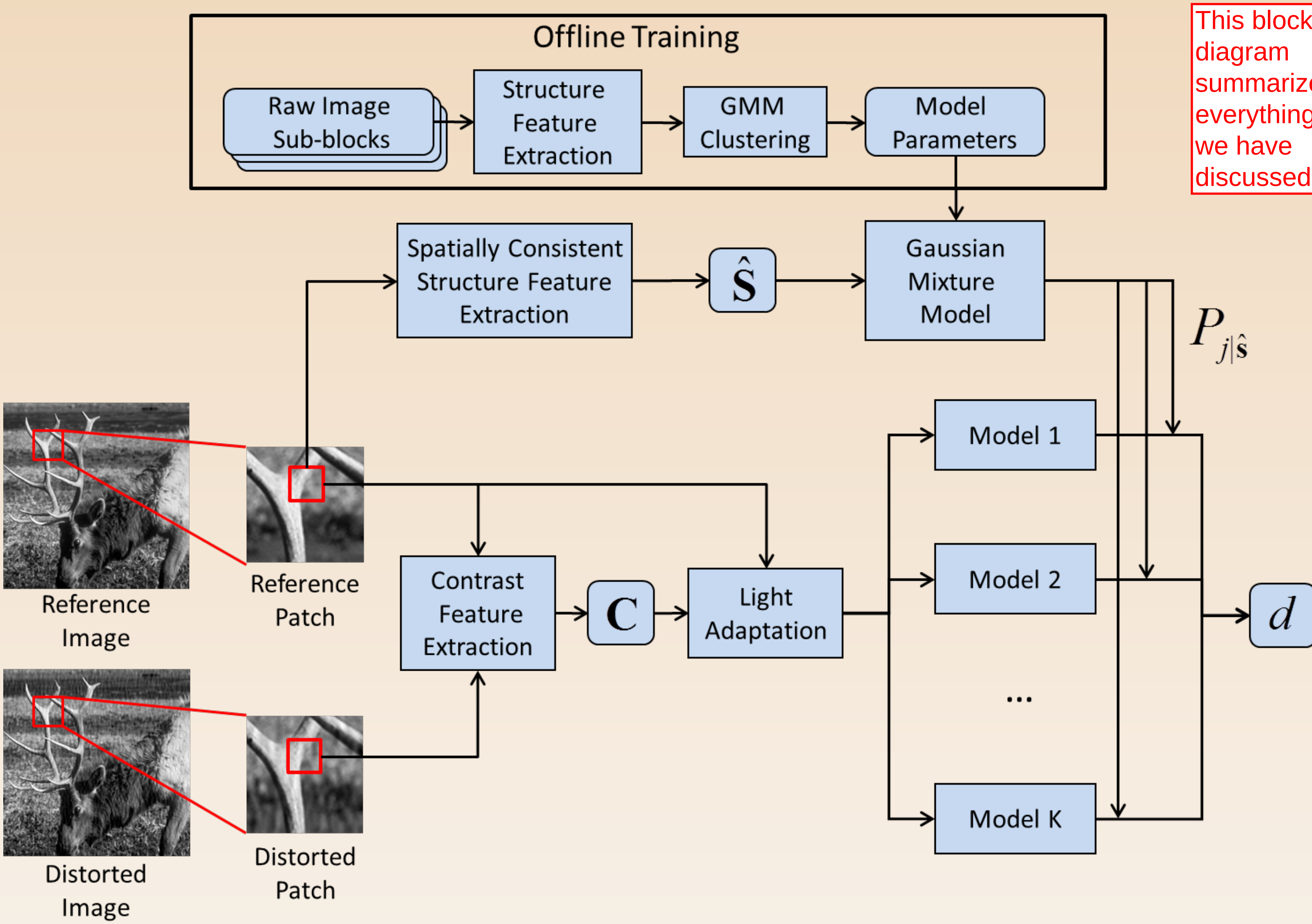
- The light adaptation function is defined using a LUT with L_b entries:

$$g(t) = g_l, \text{ where } l = \lfloor tL_b / 100 \rfloor$$

- The final distortion prediction is made by combining the model responses weighted by the structure classification probabilities:

$$d = \sum_{k=1}^{K} P(k_s = k \mid \hat{\mathbf{s}}) P(y=1 \mid k_s = k, L^*, \mathbf{c}_{err})$$
$$= \sum_{k=1}^{K} P_{k|\hat{\mathbf{s}}} \phi_k(L^*, \mathbf{c}_{err})$$

Sub-Block Distortion Prediction Framework Summary



This block diagram summarizes everything that we have discussed so far

PURDUE UNIVERSITY

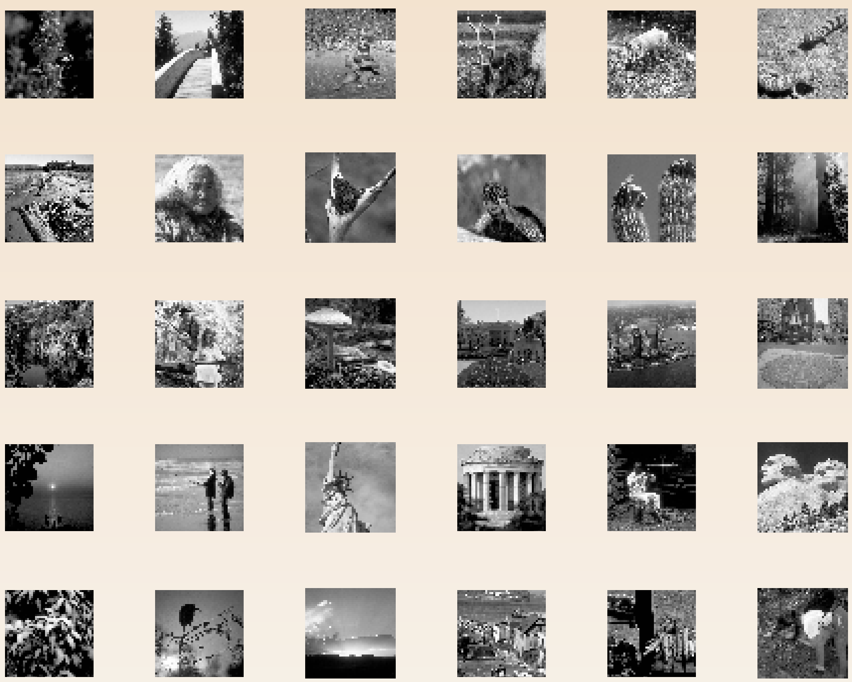
Outline

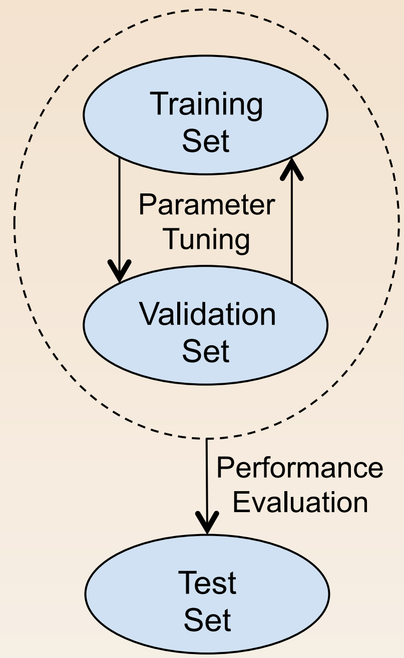
- Introduction to near-threshold local distortion in natural images
- Limitation of current methods
- Optimal classification based model for near-threshold distortion
- Results
- Conclusion

Five-Fold Cross-Validation

- We randomly divide the 1080 local patches extracted from the 30 images in the CSIQ* database into five subsets. Patches in different subsets are extracted from different images. For each local patch, the database contains error signals and contrast detection thresholds from six measurements.
- 3 subsets are used as training set, 1 subset is used as validation set, and 1 subset is used as test set.
- We perform five experiments to evaluate the model with each of the five subsets used as the test set exactly once. The average and standard deviation are recorded.

Images in CISQ database





*Alam et al, J. Vision (2014)

Md Mushfiqul Alam, Kedarnath P. Vilankar, David J. Field, and Damon M. Chandler, "Local masking in natural images: A database and analysis," J. Vision , vol. 14, no. 8, pp. 1–38, 2014.

Model Evaluation

- The ability of the model to predict near-threshold distortion is evaluated by the accuracy of distortion detection threshold prediction.
- The distortion is deemed to be present at threshold level when the model gives response of 0.5.
- The threshold prediction performance is quantified by the Pearson's Correlation Coefficient (CC), Spearman's Rank-Order Correlation Coefficient (SROCC), and RMSE between the predicted and ground-truth thresholds.
  - » Pearson's Correlation Coefficient

$$CC = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}}$$

  - » Spearman's Rank-Order Correlation Coefficient

$$\begin{array}{l} x_i \xrightarrow{rank} r_i^x \\ y_i \xrightarrow{rank} r_i^y \end{array} \Rightarrow d_i = r_i^x - r_i^y \Rightarrow SROCC = 1 - \frac{6\sum d_i^2}{n(\mathrm{n}^2-1)}$$

  - » RMSE

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2}$$

PURDUE UNIVERSITY

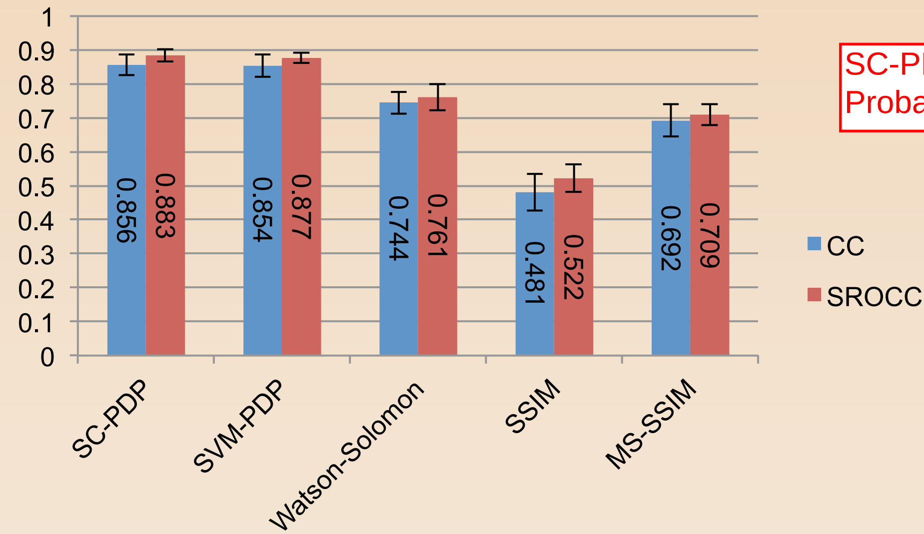
Benchmark

Structure Classification - Perceptual Distortion Prediction (SC-PDP)

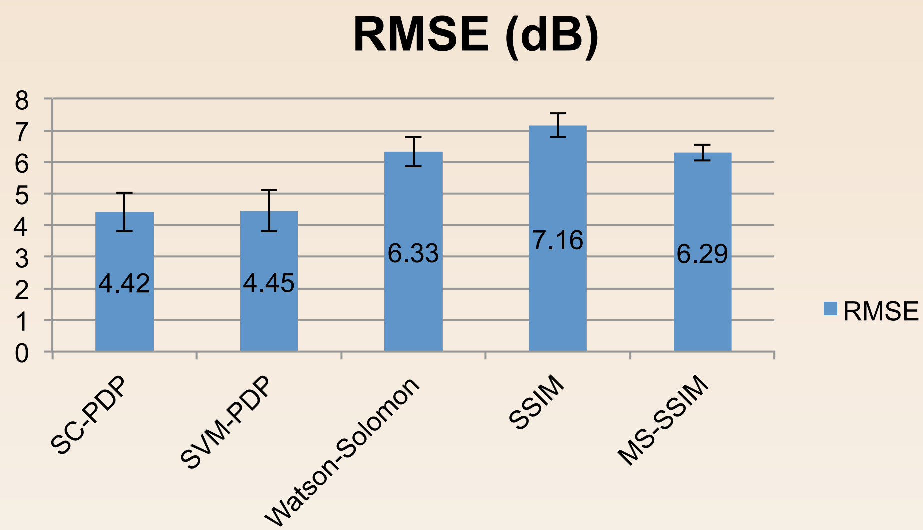
- The structure classification based distortion prediction model (SC-PDP) is compared to other perceptual similarity (distortion) models including:
  - Support vector regression using the 14 features suggested by the database owner (Alam et al.'14, very expensive and slow)
  - Watson-Solomon Model (Watson et al.'97, gain-control based model)
  - SSIM (Wang et al.'03, full-reference image quality metric)
  - Multi-scale SSIM (Wang et al.'04, full-reference image quality metric)

- All model parameters are optimized on the validation set before the performance is evaluated on the test set. (Optimal model parameter values are reported in the paper.)

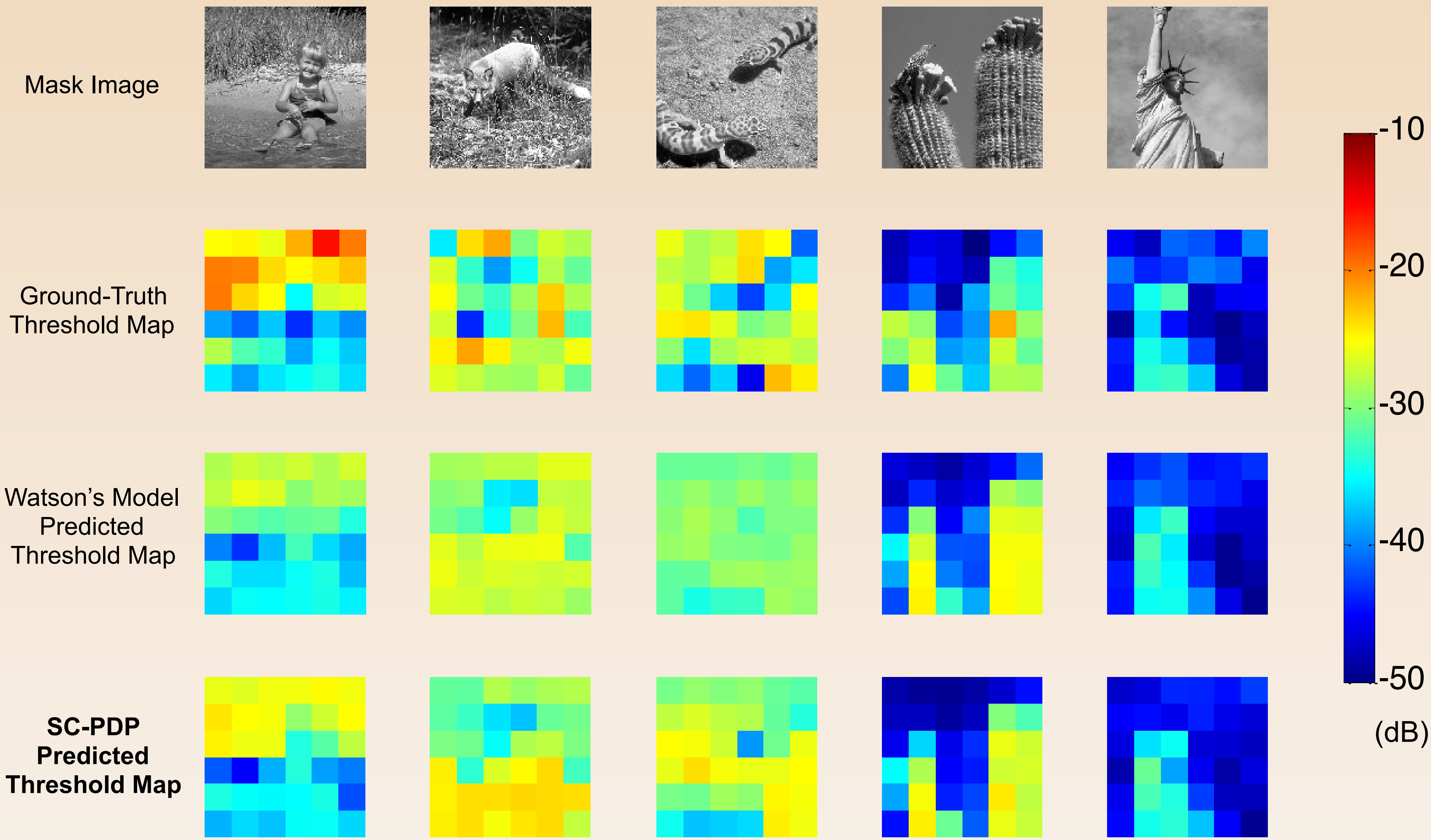
CC, SROCC and RMSE between Predicted Threshold and Ground-Truth Threshold



SC-PDP - Structure Classification - Probabilistic Distortion Predictor



Predicted Threshold Map



PURDUE UNIVERSITY

Outline

- Introduction to near-threshold local distortion in natural images
- Limitation of current methods
- Optimal classification based model for near-threshold distortion
- Results
- Conclusion

PURDUE UNIVERSITY

Conclusion

- We developed a probabilistic framework SC-PDP to better model the perception of near-threshold distortion of the HVS. The proposed model properly accounts for various aspects of the V1 response, including light adaptation, contrast sensitivity, and contrast masking of various kinds. The model achieves better threshold prediction performance than the conventional gain-control based methods and popular full-reference image quality measures.

Thank You!!