Maximum Differentiation Competition: Direct Comparison of Discriminability Models

Zhou Wang & Eero P. Simoncelli

Howard Hughes Medical Institute, Center for Neural Science, and Courant Institute for Mathematical Sciences New York University





## Which model best accounts for perceived image quality?



## Which model best accounts for perceived image quality?

#### **Example Models**

MSE: Mean Squared Error  $E(\mathbf{X}, \mathbf{Y}) = \frac{1}{N} \sum_{i} (x_i - y_i)^2$ 

SSIM: Structural Similarity [Wang, et. al. '04]

- local cross-correlation measure:

$$s(\mathbf{x}, \mathbf{y}) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

- pooling 
$$S(\mathbf{X}, \mathbf{Y}) = \frac{\sum_{i} w(\mathbf{x}_{i}, \mathbf{y}_{i}) s(\mathbf{x}_{i}, \mathbf{y}_{i})}{\sum_{i} w(\mathbf{x}_{i}, \mathbf{y}_{i})}$$

where  $w(\mathbf{x}, \mathbf{y}) = \log_2(1 + \sigma_x^2/C) + \log_2(1 + \sigma_y^2/C)$ 

Wang & Simoncelli, VSS-2005

## **Conventional Method**

- Procedure
  - 1. Choose set of reference and distorted images
  - 2. Perform subjective tests
  - 3. Compare model prediction with subjective responses

- Difficulties
  - Subjective experiments expensive
  - "Curse of dimensionality": impossible to cover image space

#### **Conventional Method: MSE vs. SSIM**



#### "LIVE" image database, UT Austin

#### **Conventional Method: MSE vs. SSIM**



| Distortion: | JP2(1) | JP2(2) | JPG(1) | JPG(2) | Noise | Blur  | Error |
|-------------|--------|--------|--------|--------|-------|-------|-------|
| # images:   | 87     | 82     | 87     | 88     | 145   | 145   | 145   |
| MSE         | 0.934  | 0.895  | 0.902  | 0.914  | 0.987 | 0.774 | 0.881 |
| SSIM        | 0.968  | 0.967  | 0.965  | 0.986  | 0.971 | 0.936 | 0.944 |

## **MAximum Differentiation (MAD) Competition**

## **MAximum Differentiation (MAD) Competition**

• Let two models **compete** 

## **MAximum Differentiation (MAD) Competition**

- Let two models **compete**
- ... by **synthesizing** optimal stimuli

## **MAximum Differentiation (MAD) Competition**

- Let two models **compete**
- ... by **synthesizing** optimal stimuli
- ... that **maximally differentiate** the models

























# **2AFC Experiment**

#### distortion level (MSE)



# **2AFC Experiment**

#### distortion level (MSE)



- Subjects: 5 (4 naïve, 1 author)
- Images: 10 reference, viewed at 16 pixels/degree
- Trials: 20 per distortion-level per subject

#### **Psychometric Functions**



initial distortion level (MSE)

## **Psychometric Functions**



## Summary

- MAximum Differentiation (MAD) Competition
  - Let two models compete
  - ... by synthesizing optimal stimuli
  - ... that maximally differentiate the models
- Advantages
  - Optimized images maximize opportunity for model failure
  - Efficient (minimal # of 2-alternative comparisons)
  - Images reveal model weaknesses => potential improvements
- To Do
  - Full experiment, with more reference images
  - Application to other discriminable quantities
  - Physiology