GROUP MAD COMPETITION - A NEW METHODOLOGY TO COMPARE OBJECTIVE IMAGE QUALITY MODELS

Kede Ma, **Qingbo Wu**, Zhou Wang, Zhengfang Duanmu,
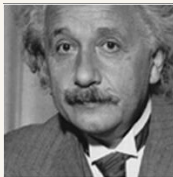Hongwei Yong, Hongliang Li and Lei Zhang

June 28, 2016

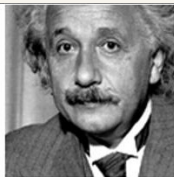# Image Quality Assessment (IQA)

Purpose

Create objective models to predict human perception of image quality.
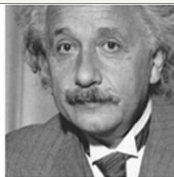
Question

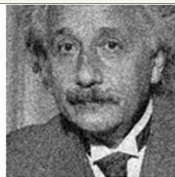With a significant number of IQA models available, how to fairly compare their performance?

(a) MSE=0, SSIM=1
CW-SSIM=1

(b) MSE=306, SSIM=0.928
CW-SSIM=0.938

(c) MSE=309, SSIM=0.987
CW-SSIM=1.000

(d) MSE=309, SSIM=0.576
CW-SSIM=0.814

Conventional Evaluation Methodology

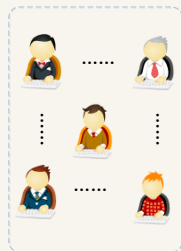Prove them by computing correlation metrics between subjective assessment and objective model predictions.

Problem

Enormous image space
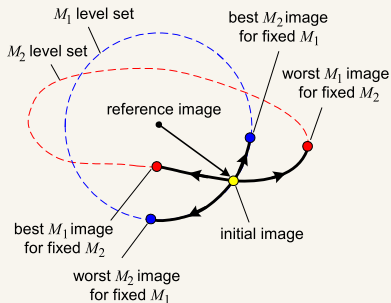

Unaffordable

Subjective test

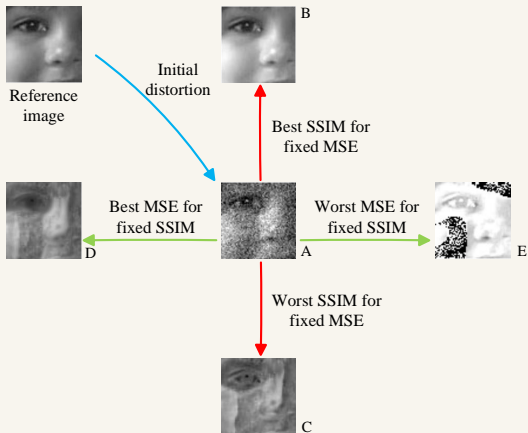# MAximum Differentiation (MAD) Competition

Merits of MAD

Disprove IQA models by synthesizing strongest "counter-examples".

Counter-examples search

M_1 level set

M_2 level set

best M_2 image for fixed M_1

worst M_1 image for fixed M_2

reference image

best M_1 image for fixed M_2

worst M_2 image for fixed M_1

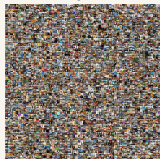initial image

MAD Competition
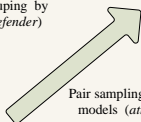
Attacking-Defending Game between Models

(a) An image collection

image grouping by
PSNR (*defender*)

(b) Subset of images that
have the same PSNR

Pair sampling by other
models (*attackers*)

MS-SSIM
(*attacher1*)

(c1) Best MS-SSIM image

(c2) Worst MS-SSIM image

BIQI
(*attacher2*)

(d1) Best BIQI image

(d2) Worst BIQI image

M3
(*attacher3*)

(e1) Best M3 image

(e2) Worst M3 image

Subjective Testing

Group MAD (gMAD) Competition

Performance Measures

1. Aggressiveness: How successful of a model at attacking another model?
2. Resistance: How successful of a model at defending the attacks from another model?
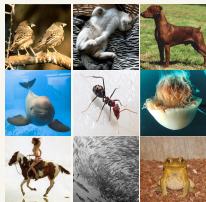
Global Ranking

The Global rankings obtained by aggregating the *aggressiveness* matrix **A** and *Resistance* matrix **R**.
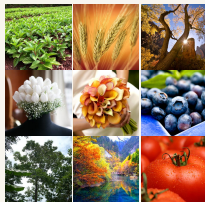
Waterloo Exploration Database

4K+ source and ∼100K distorted images

Human

Animal
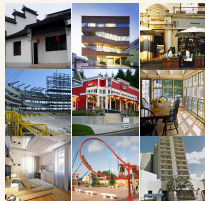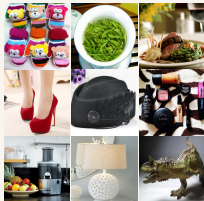

Plant

Landscape

Cityscape

Still-life

Transportation

Pairwise Comparison between 16 Models
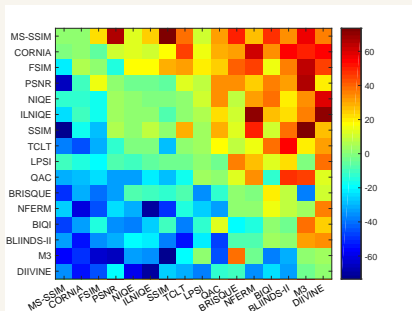
Figure: *Aggressiveness* matrix
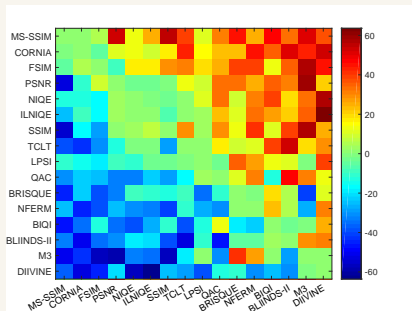
Figure: *Resistance* matrix

Global Ranking Result of 16 Models
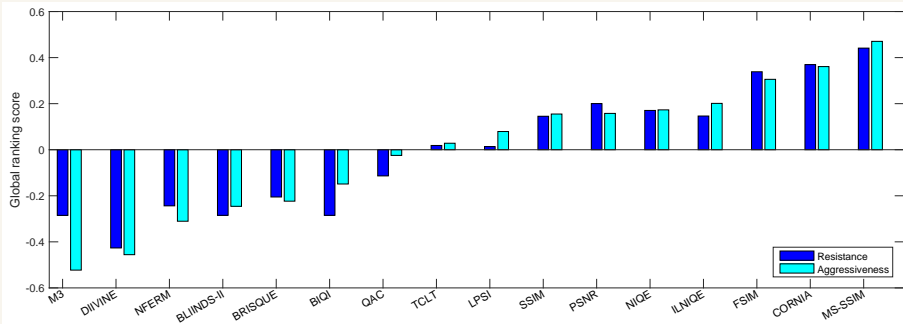
Observations

1. FR-IQA models are more competitive;
2. MS-SSIM and FSIM are top performing FR-IQA models;
3. CORNIA and ILNIQE are top performing NR-IQA models;
4. Machine learning based IQA models generally do not perform well.

Thank you