Image Quality Assessment: From Error Visibility to Structural Similarity

Zhou Wang

Motivation

original Image

MSE=0, MSSIM=1

MSE=225, MSSIM=0.949

MSE=225, MSSIM=0.989

MSE=215, MSSIM=0.671

MSE=225, MSSIM=0.688

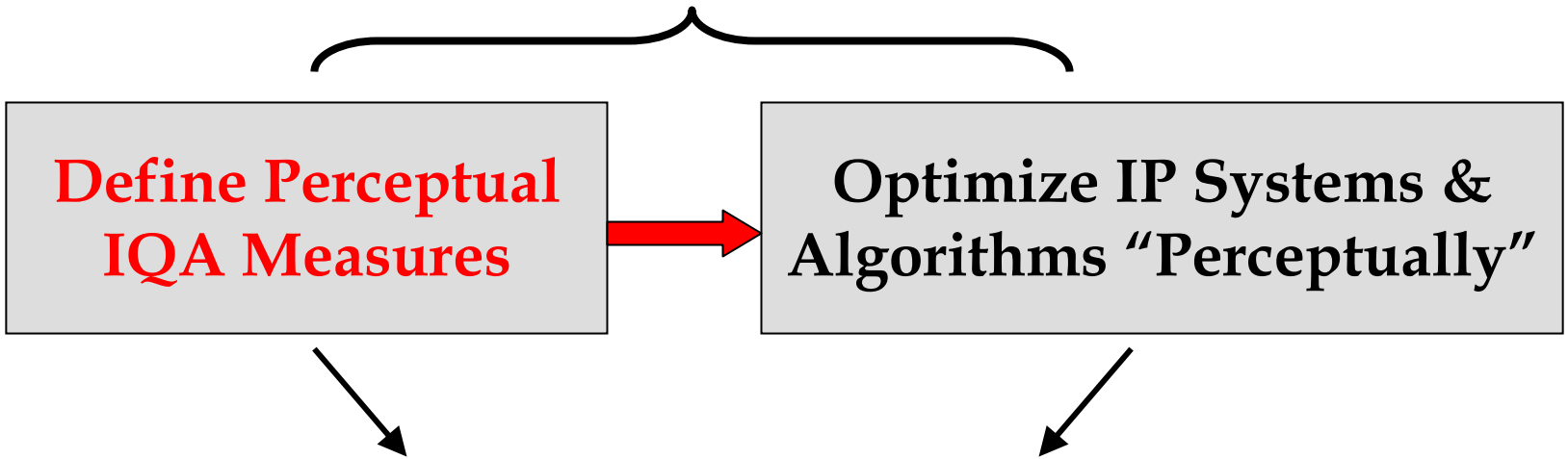
MSE=225, MSSIM=0.723

Perceptual Image Processing

Why? Standard measure (MSE) does not agree with human visual perception

↓

PERCEPTUAL IMAGE PROCESSING



Application Scope: essentially all IP applications

image/video compression, restoration, enhancement, watermarking, displaying, printing …

Image Quality Assessment

- Goal
  - Automatically predict perceived image quality
- Classification
  - Full-reference (FR); No-reference (NR); Reduced-reference (RR)
- Widely Used Methods
  - FR: MSE and PSNR → $PSNR = 10\log_{10}\frac{L^2}{MSE}$
  - NR & RR: wide open research topic
- **IQA is Difficult**

VQEG (1)

- VQEG (video quality experts group)
  1. Goal: recommend video quality assessment standards (TV, telecommunication, multimedia industries)
  2. Hundreds of experts (*Intel, Philips, Sarnoff, Tektronix, AT&T, NHK, NASA, Mitsubishi, NTIA, NIST, Nortel ……*)

- Testing methodology
  1. Provide test video sequences
  2. Subjective evaluation
  3. Objective evaluation by VQEG proponents
  4. Compare subjective/objective results, find winner

VQEG (2)

- Current Status
  1. Phase I test (2000):
     - Diverse types of distortions
     - 10 proponents including PSNR
     - no winner, 8~9 proponents statistically equivalent, including **PSNR!**

  2. Phase II test (2003):
     - Restricted types of distortions (MPEG)
     - Result: A few models slightly better than PSNR

  3. VQEG is extending their directions:
     - FR/RR/NR, Low Bit Rate
     - Multimedia: video, audio and speech …

Standard IQA Model: Error Visibility (1)

Philosophy

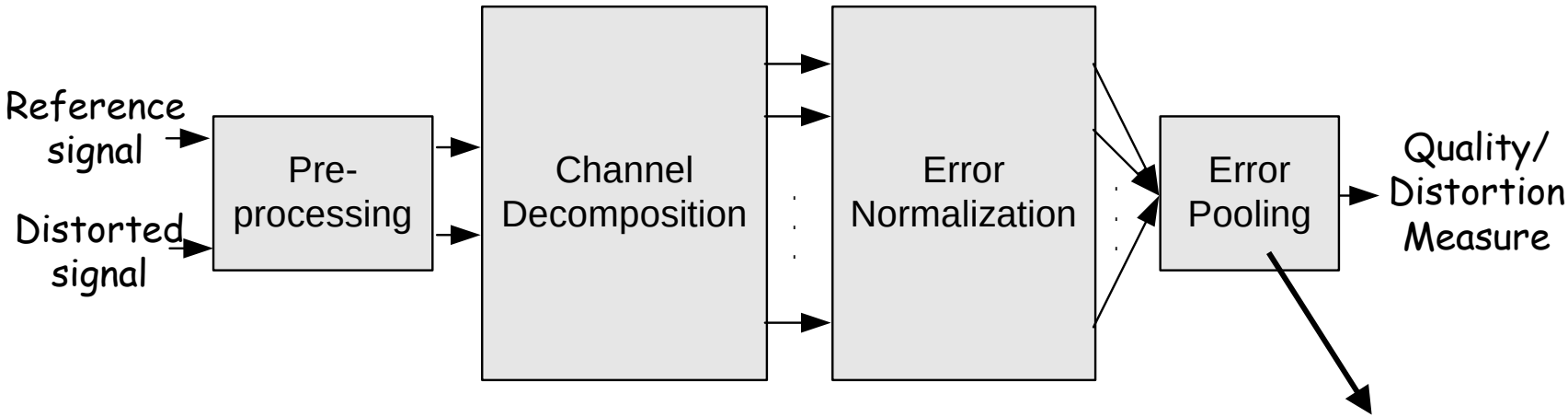
distorted signal = reference signal + error signal

Assume reference signal has perfect quality

Quantify **perceptual** error visibility

- **Representative work**
  - Pioneering work [Mannos & Sakrison ’74]
  - Sarnoff model [Lubin ’93]
  - Visible difference predictor [Daly ’93]
  - Perceptual image distortion [Teo & Heeger ’94]
  - DCT-based method [Watson ’93]
  - Wavelet-based method [Safranek ’89, Watson *et al.* ’97]

Standard IQA Model: Error Visibility (2)



- Motivation

  Simulate relevant early HVS components

$$E = \left[\sum_{l} \sum_{k} \left| e_{l,k} \right|^{\beta} \right]^{1/\beta}$$

- Key features

  Channel decomposition → linear frequency/orientation transforms

  Frequency weighting → contrast sensitivity function

  Masking → intra/inter channel interaction

Standard IQA Model: Error Visibility (3)

- Quality **definition** problem
  - Error visibility = quality ?
- The **suprathreshold** problem
  - Based on threshold psychophysics
  - Generalize to suprathreshold range?
- The natural image **complexity** problem
  - Based on simple-pattern psychophysics
  - Generalize to complex natural images?

[Wang, *et al*., "Why is image quality assessment so difficult?" *ICASSP* '02]
[Wang, *et al*., *IEEE Trans. Image Processing,* '04]

New Paradigm: Structural Similarity

Philosophy

Purpose of human vision: extract **structural information**

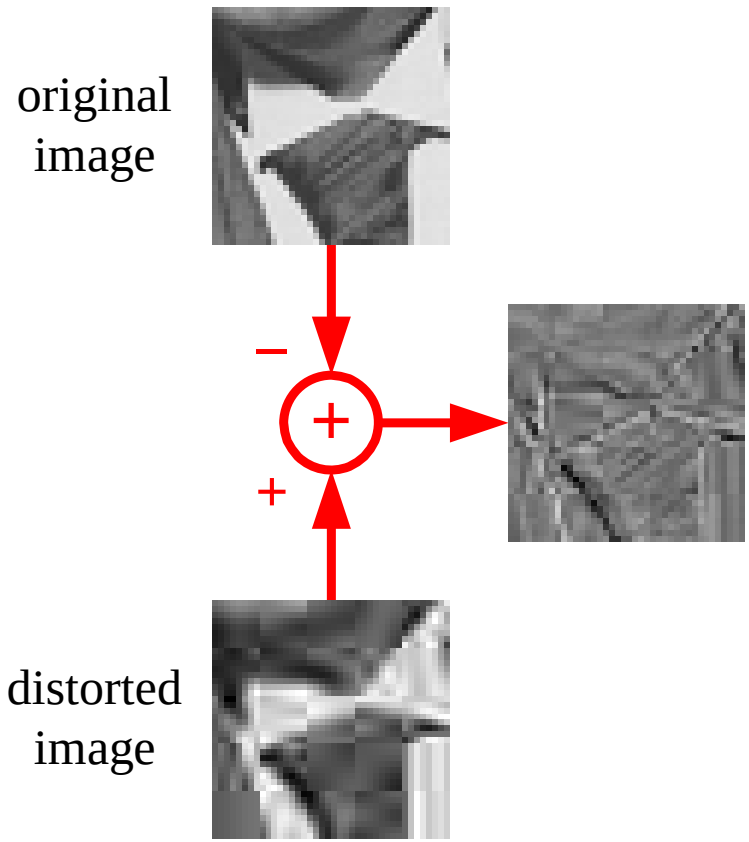
HVS is highly adapted for this purpose

Estimate **structural information change**

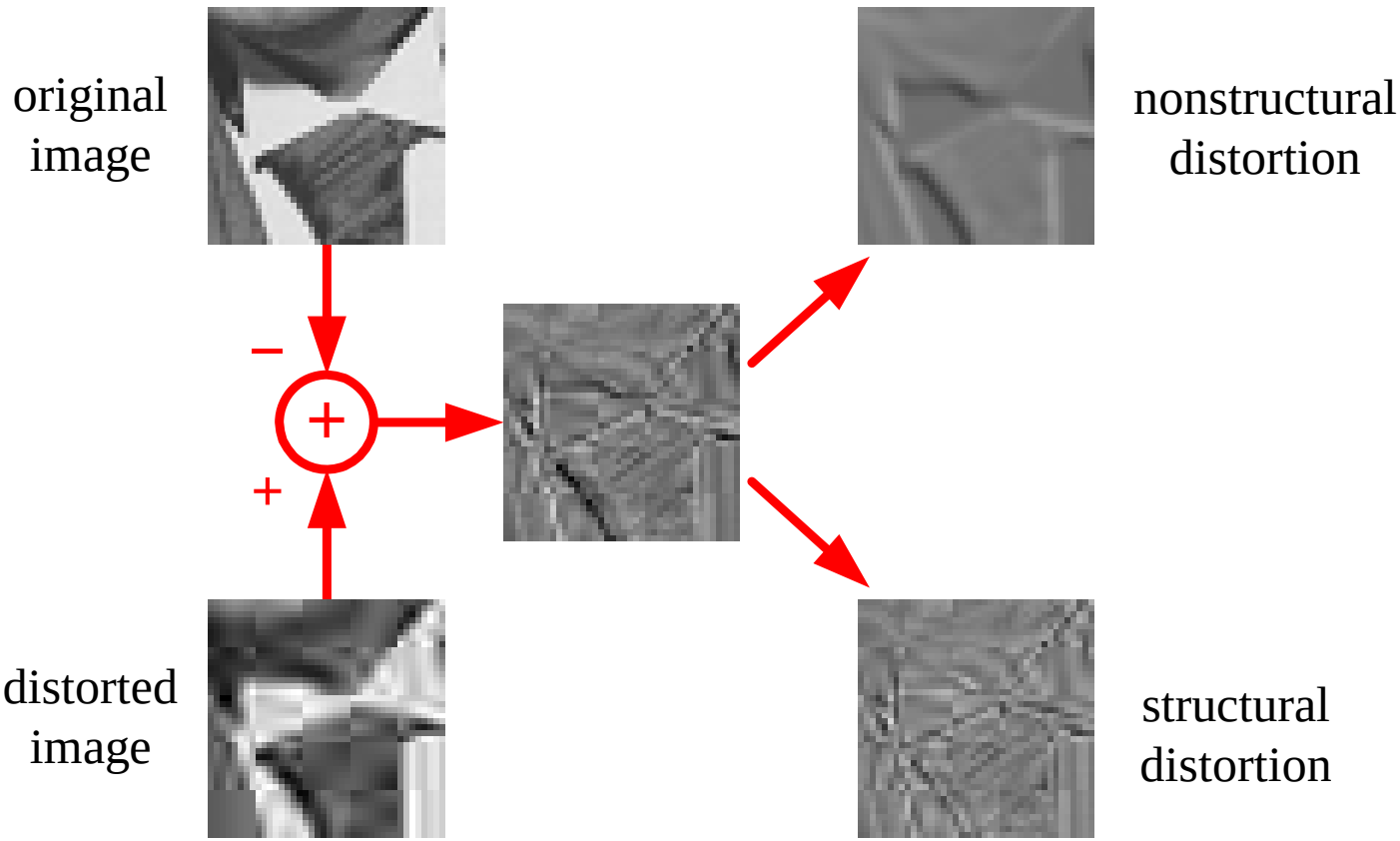
Classical philosophy	New philosophy
Bottom-up	Top-down
Predict Error Visibility	Predict Structural Distortion

- How to define structural information?
- How to separate structural/nonstructural information?

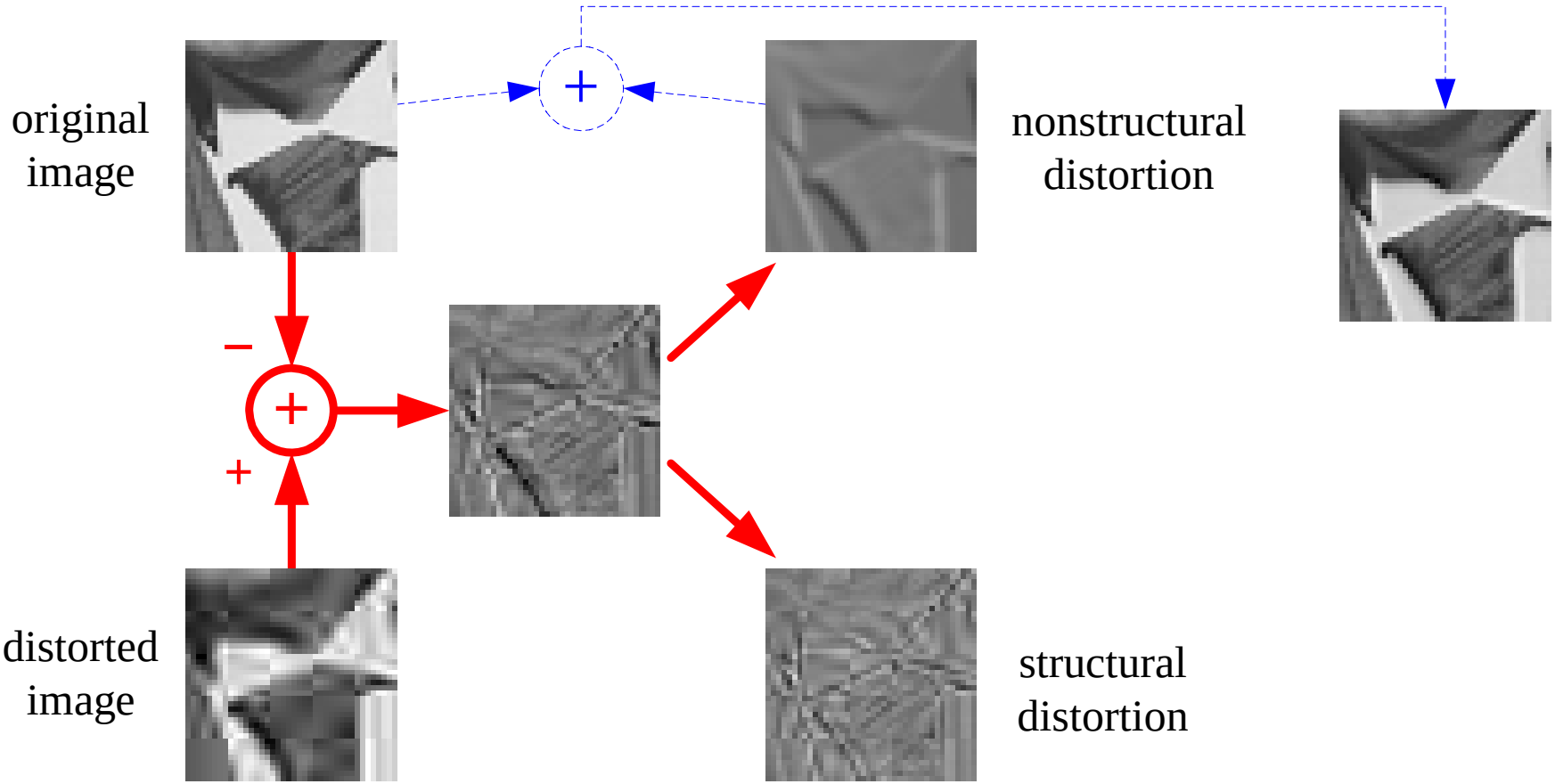
Separation of Structural/nonstructural Distortion



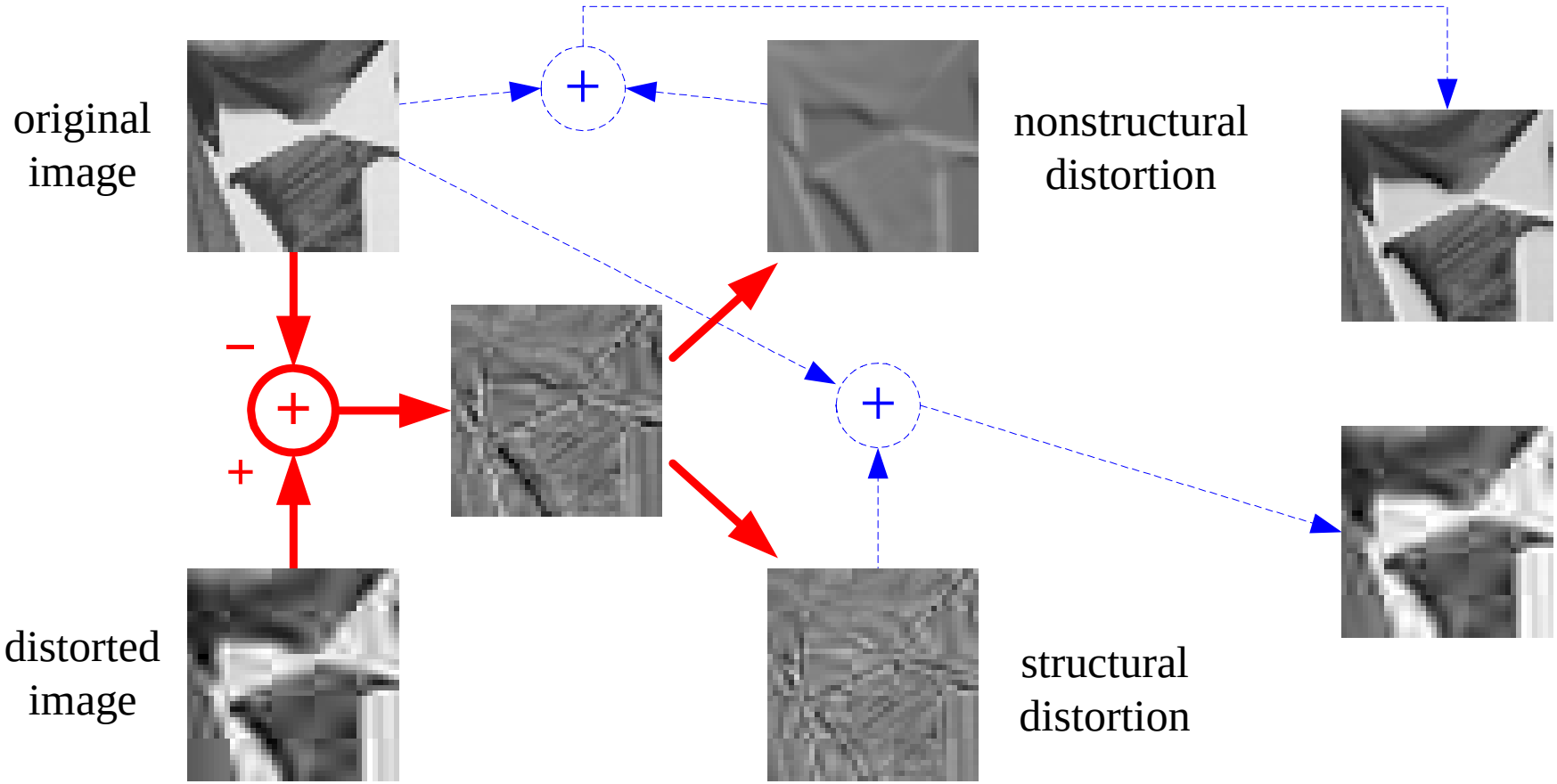
Separation of Structural/nonstructural Distortion



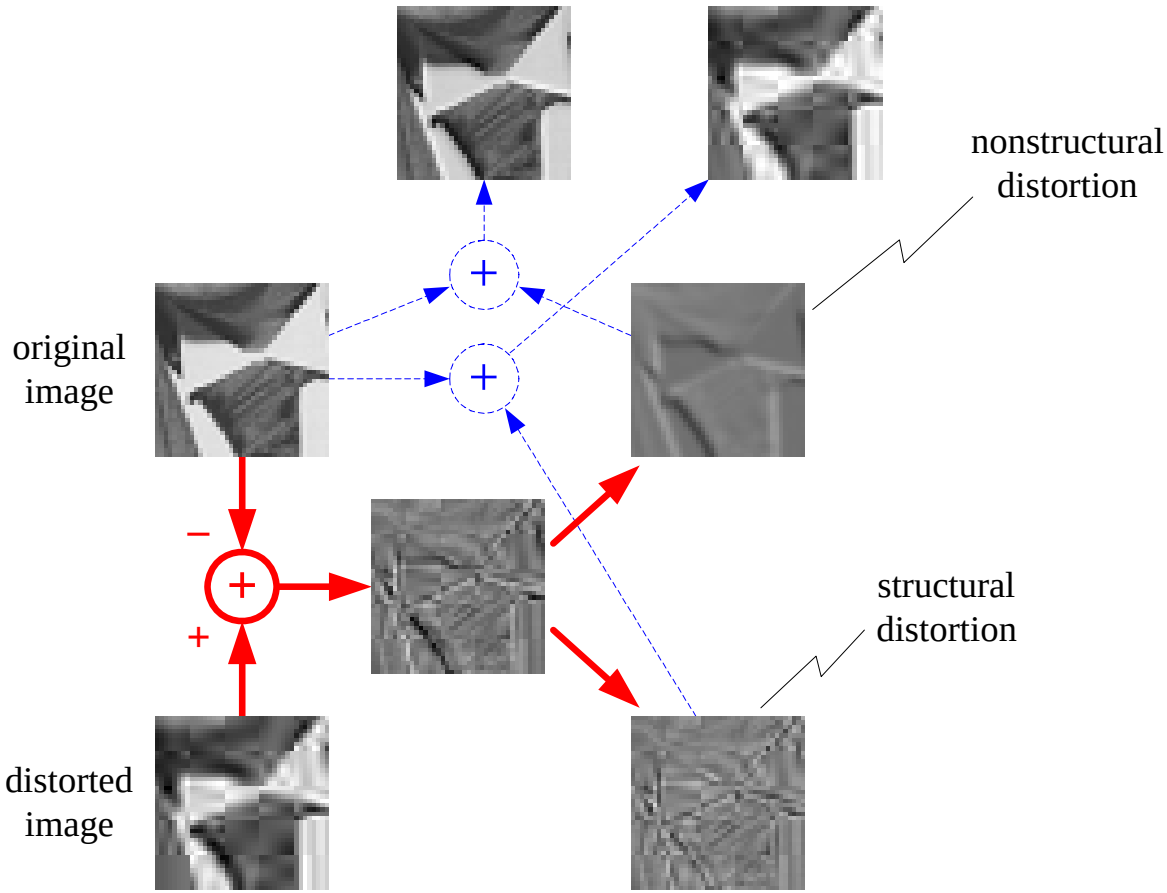
Separation of Structural/nonstructural Distortion



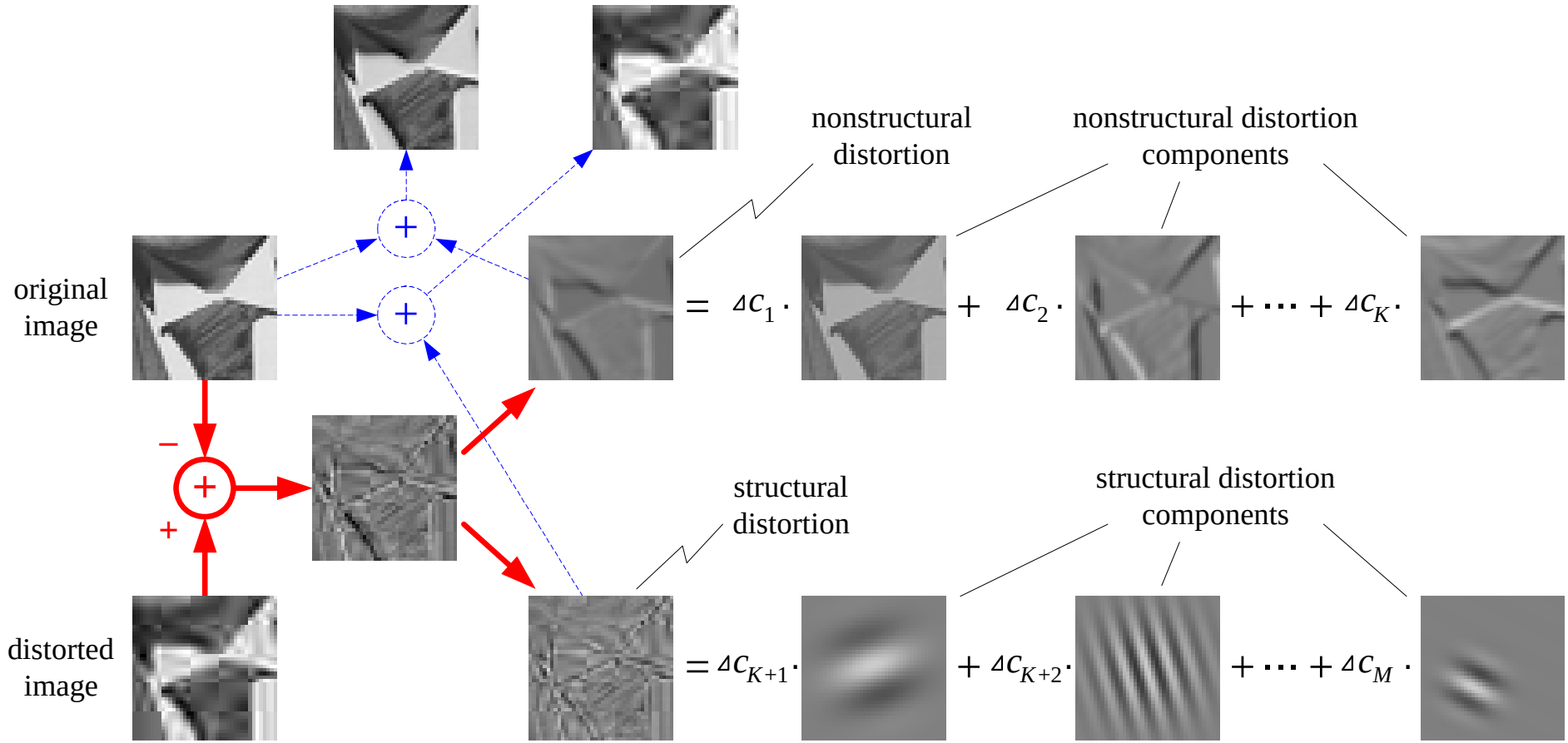
Separation of Structural/nonstructural Distortion



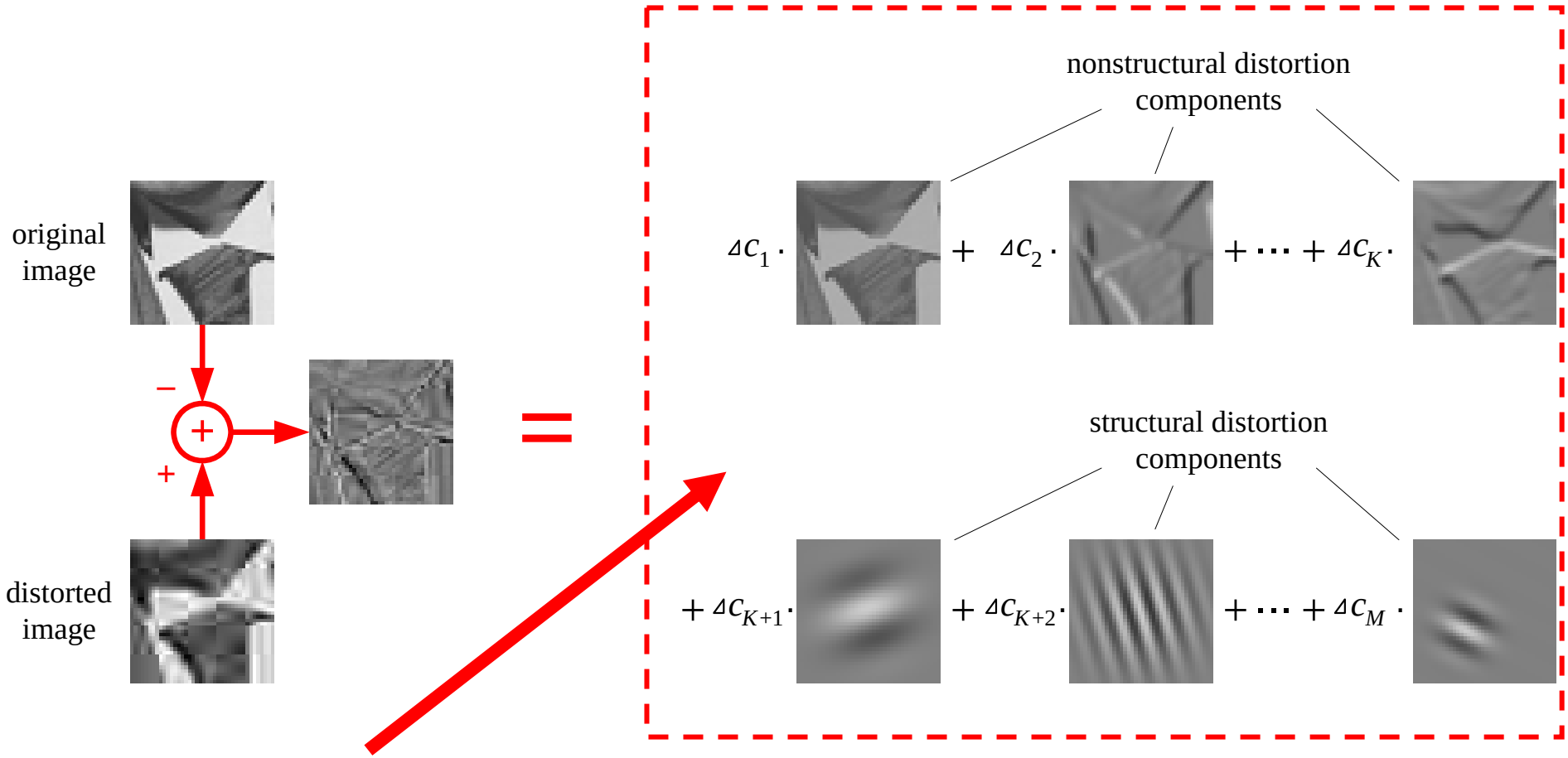
Adaptive Linear System



Adaptive Linear System



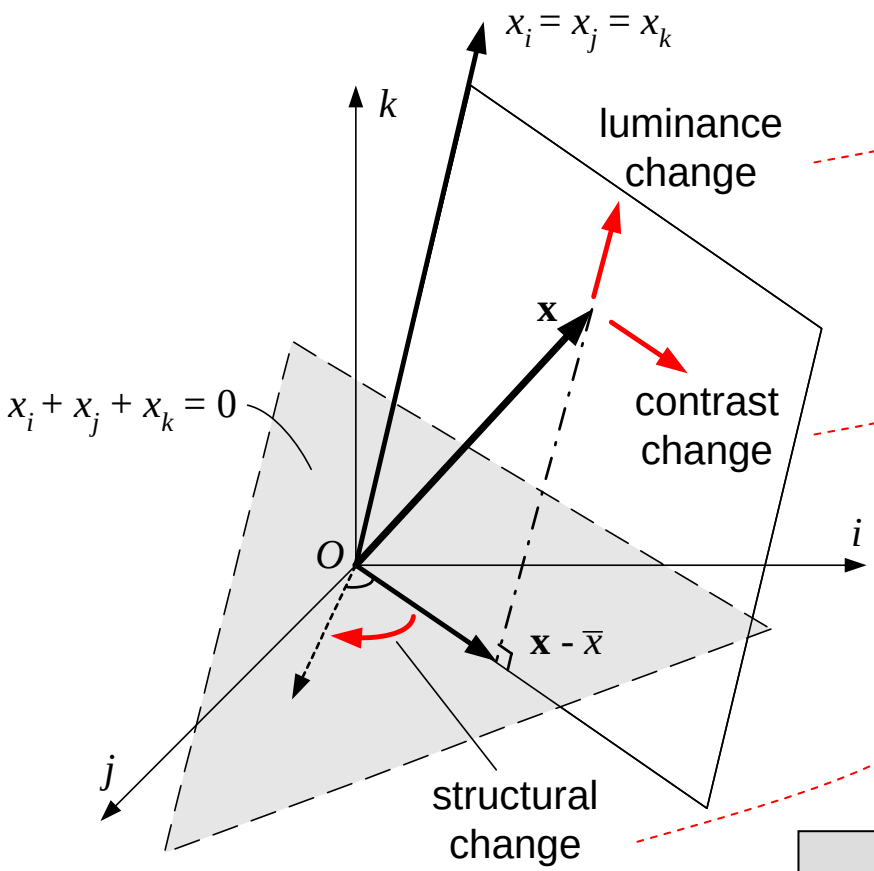
Adaptive Linear System



overcomplete, adaptive basis in the space of all images

[Wang & Simoncelli, *ICIP* '05, submitted]

Structural Similarity (SSIM) Index in Image Space



$$l(\mathbf{x},\mathbf{y}) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}$$

$$c(\mathbf{x},\mathbf{y}) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}$$

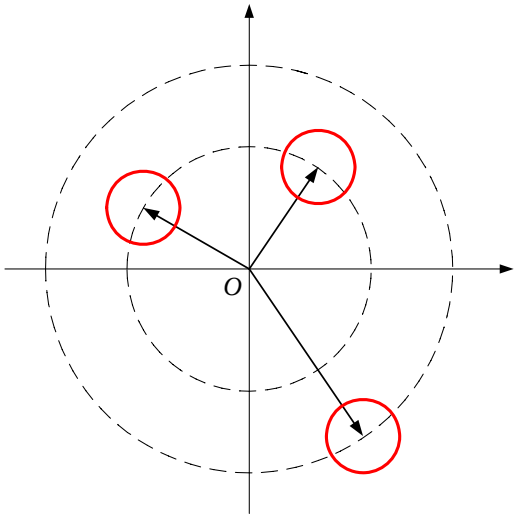
$$s(\mathbf{x},\mathbf{y}) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}$$

$$SSIM(\mathbf{x},\mathbf{y}) = l(\mathbf{x},\mathbf{y}) \cdot c(\mathbf{x},\mathbf{y}) \cdot s(\mathbf{x},\mathbf{y})$$

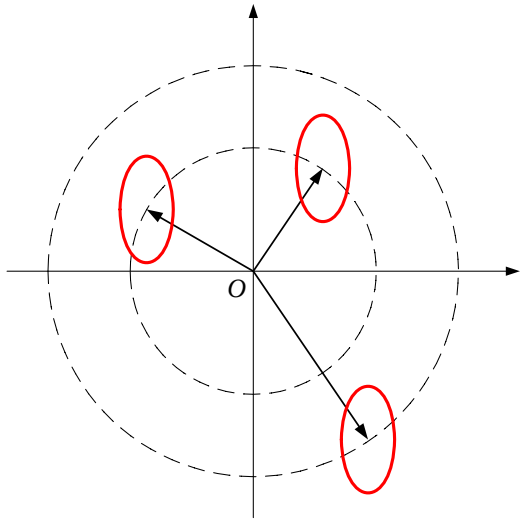
[Wang & Bovik, *IEEE Signal Processing Letters*, '02]

[Wang *et al*., *IEEE Trans. Image Processing,* '04]

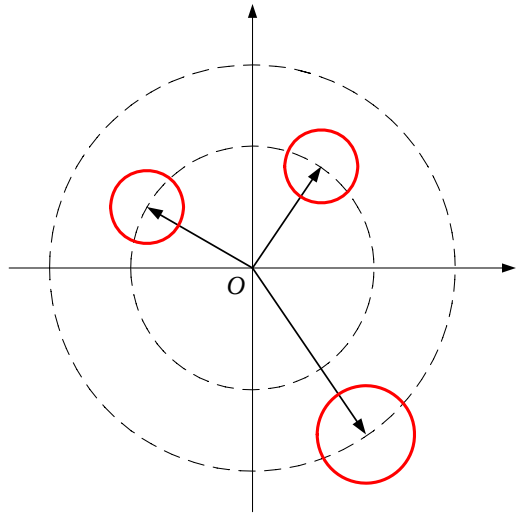
Model Comparison



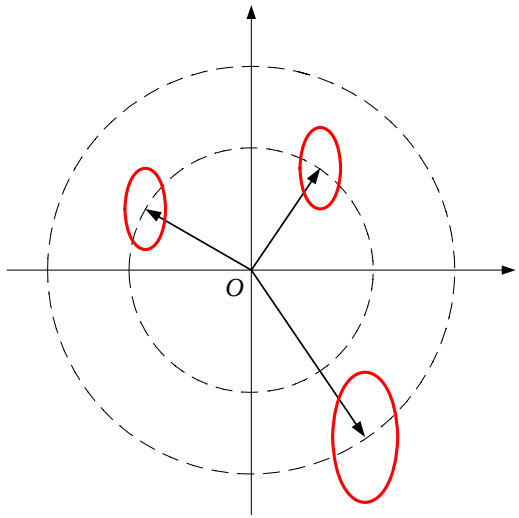
Minkowski (MSE)



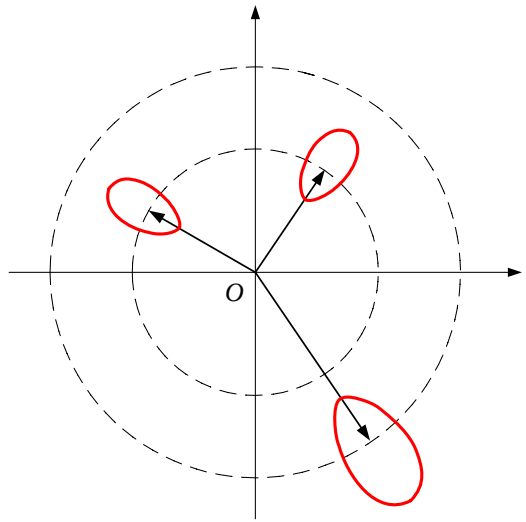
component-weighted



magnitude-weighted



magnitude and component-weighted



SSIM

JPEG2000 compressed image

original image

SSIM index map

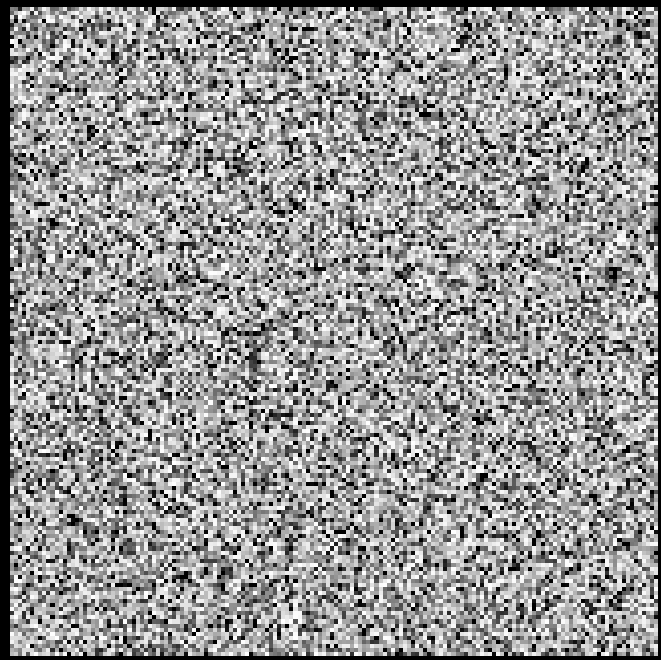
absolute error map

Gaussian noise corrupted image

original image

SSIM index map

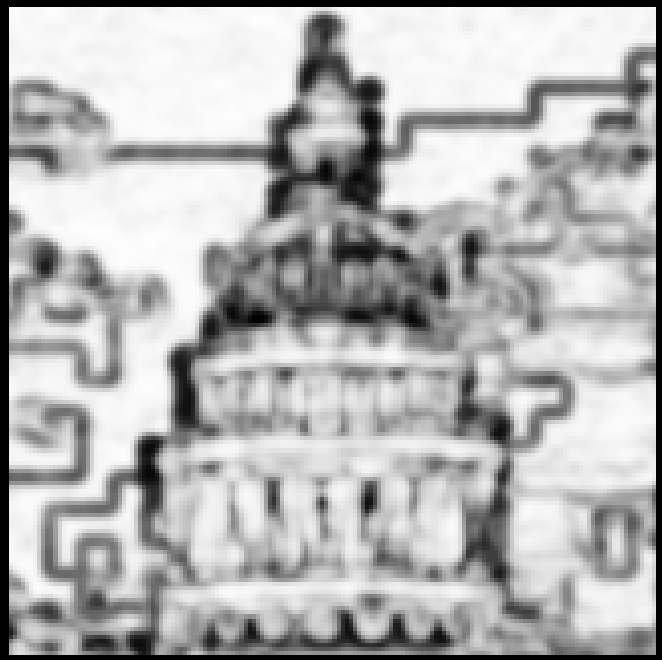
absolute error map



JPEG compressed image

original image

SSIM index map



absolute error map

Demo Images

MSE=0, MSSIM=1

MSE=225, MSSIM=0.949

MSE=225, MSSIM=0.989

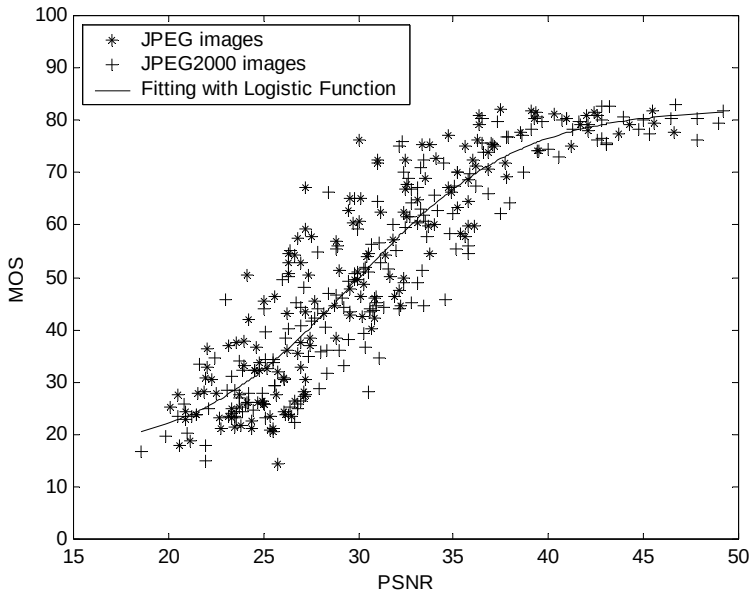
MSE=215, MSSIM=0.671

MSE=225, MSSIM=0.688

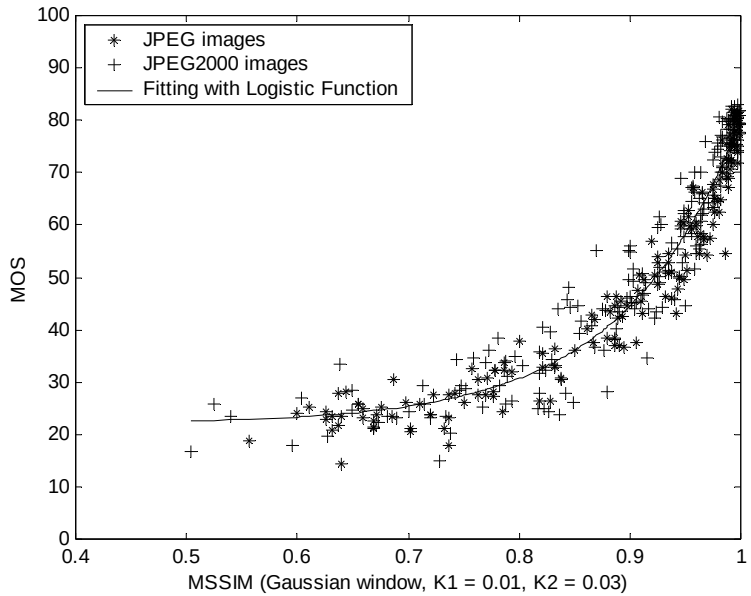
MSE=225, MSSIM=0.723

Validation LIVE Database

Dataset	JP2(1)	JP2(2)	JPG(1)	JPG(2)	Noise	Blur	Error
# of images	87	82	87	88	145	145	145
PSNR	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**

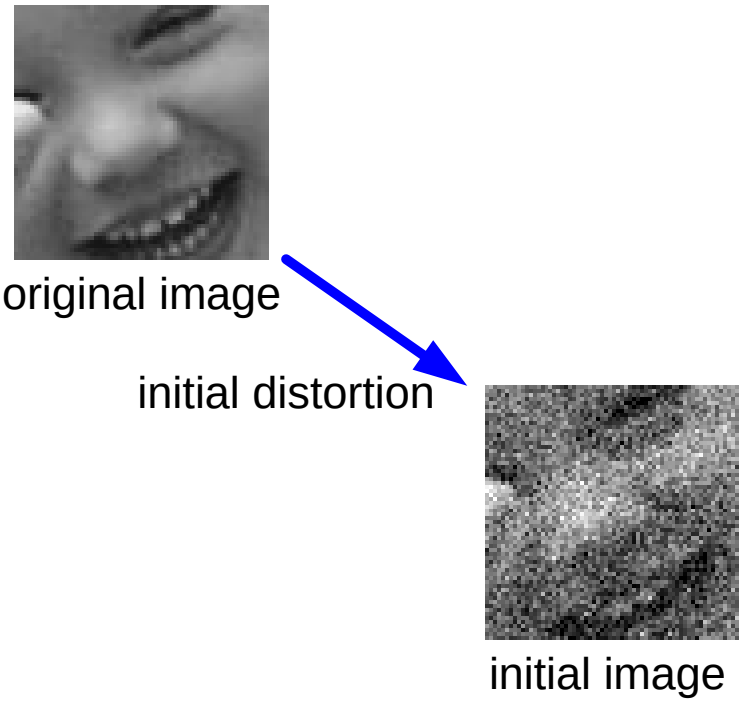


PSNR



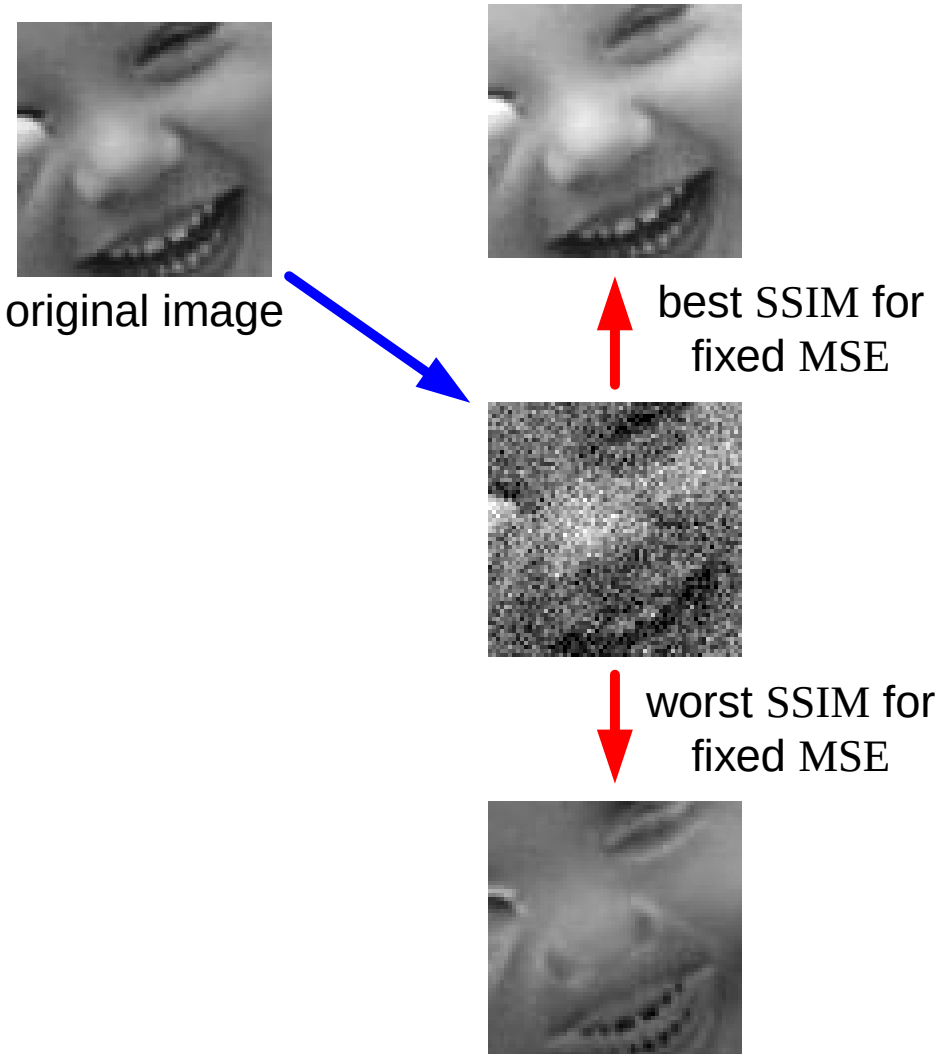
MSSIM

MAD Competition: MSE vs. SSIM (1)



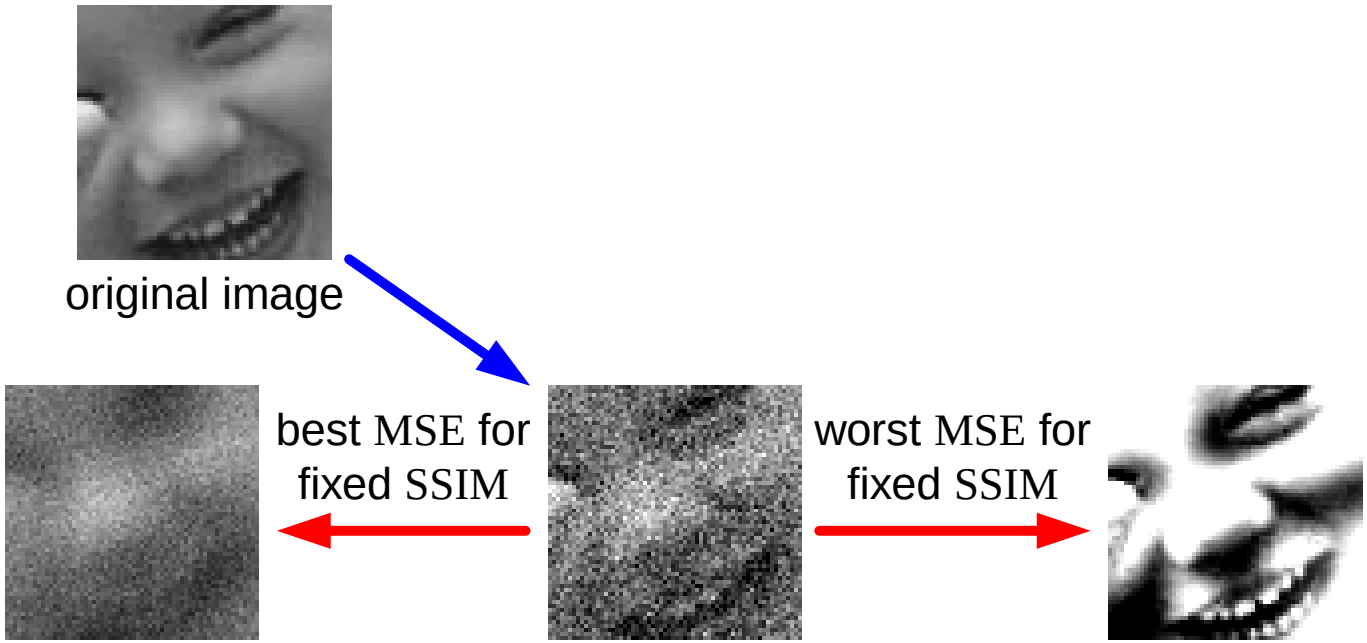
[Wang & Simoncelli, *Human Vision and Electronic Imaging,* ’04]

MAD Competition: MSE vs. SSIM (2)



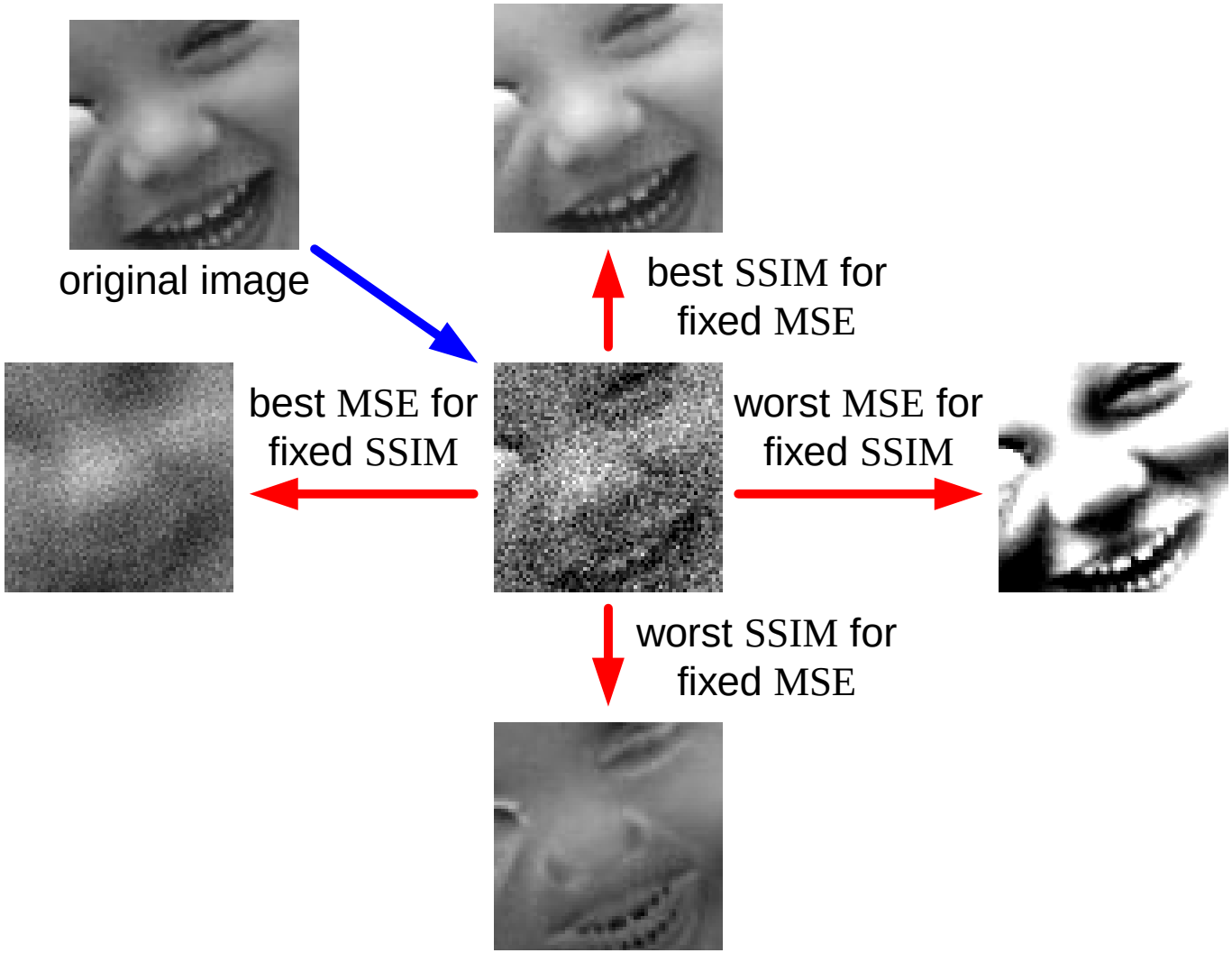
[Wang & Simoncelli, *Human Vision and Electronic Imaging,* ’04]

MAD Competition: MSE vs. SSIM (3)



[Wang & Simoncelli, *Human Vision and Electronic Imaging,* '04]

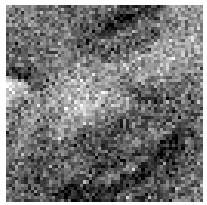
MAD Competition: MSE vs. SSIM (4)



[Wang & Simoncelli, *Human Vision and Electronic Imaging,* '04]

original image

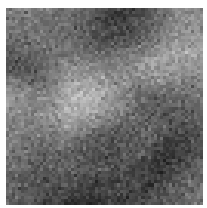
initial distorted image



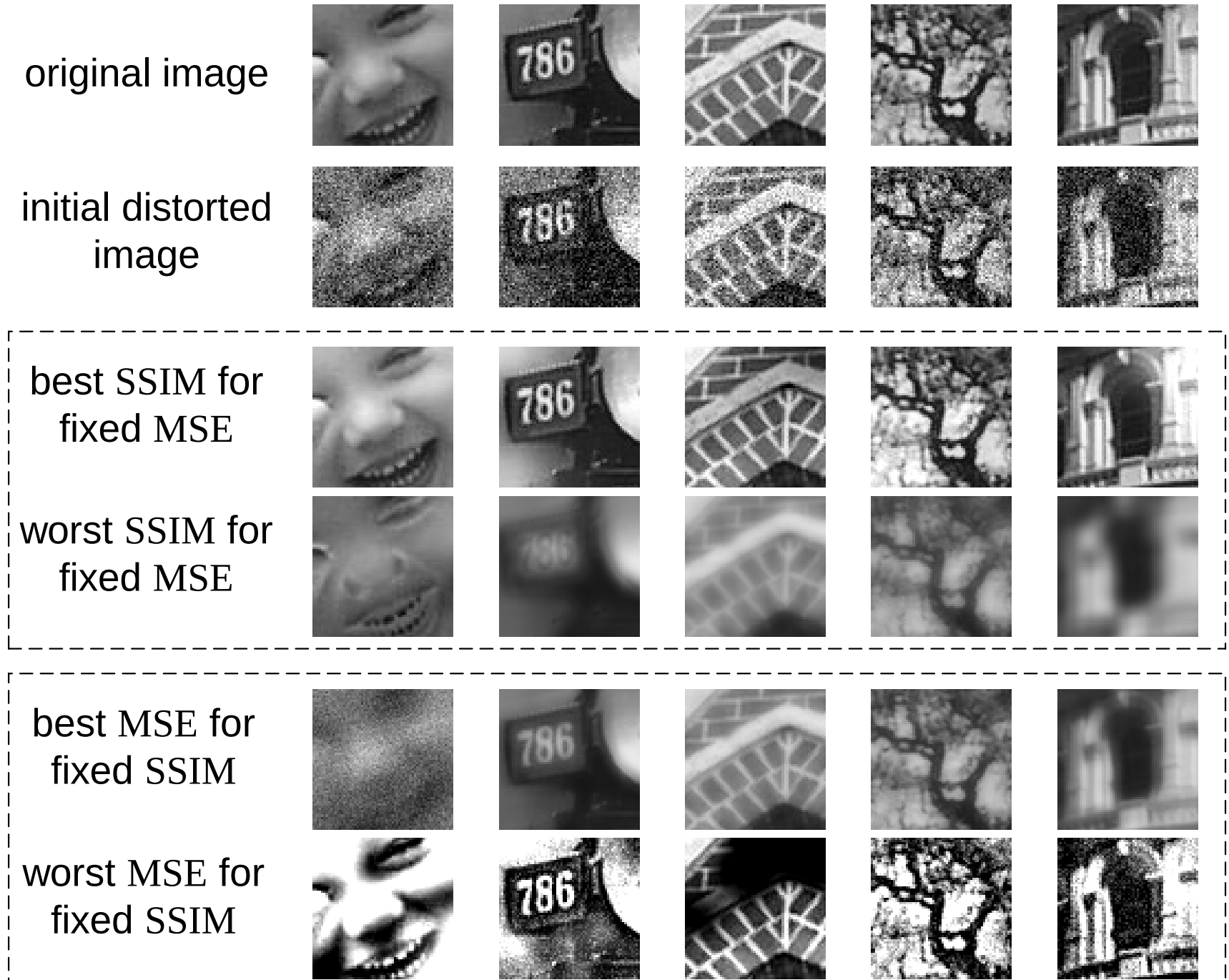
best SSIM for fixed MSE

worst SSIM for fixed MSE

best MSE for fixed SSIM



worst MSE for fixed SSIM


original image
initial distorted image
best SSIM for fixed MSE
worst SSIM for fixed MSE
best MSE for fixed SSIM
worst MSE for fixed SSIM

Extensions of SSIM (1)

- **Color image quality assessment**

  [Toet & Lucassen., *Displays,* ’03]

- **Video quality assessment**

  [Wang, *et al*., *Signal Processing: Image Communication,* ’04]

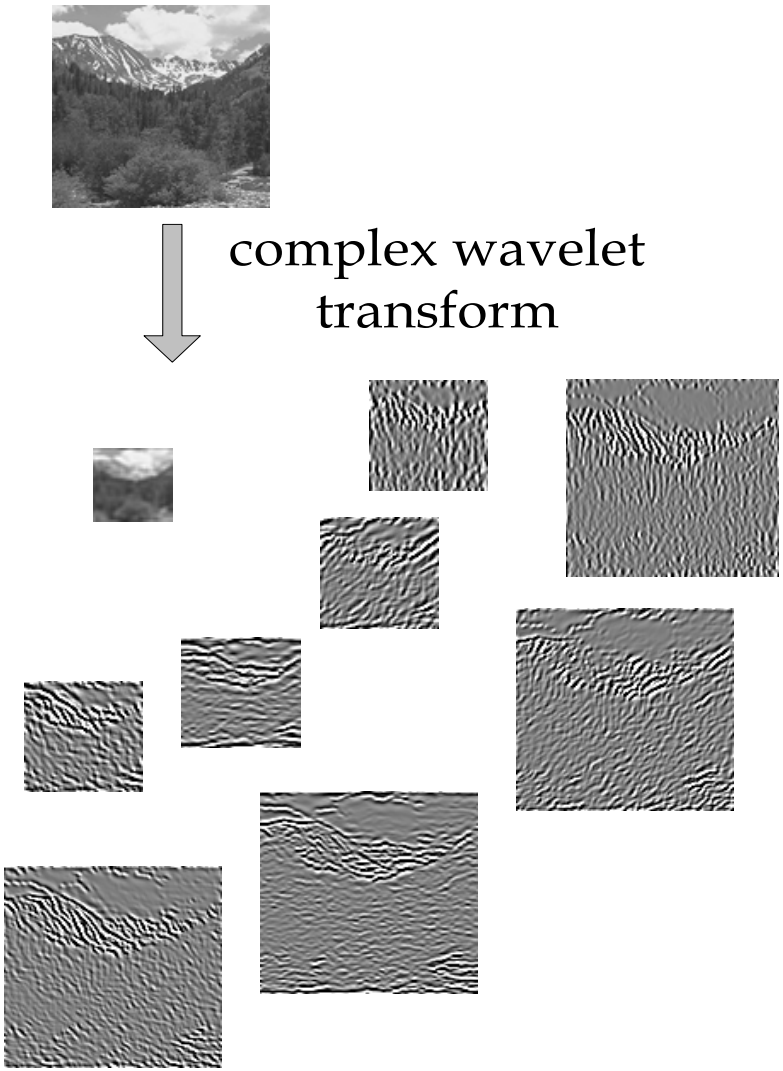
- **Multi-scale SSIM**

  [Wang, *et al*., Invited Paper, *IEEE Asilomar Conf.* ’03]

- **Complex wavelet SSIM**

  [Wang & Simoncelli, *ICASSP* ’05]

Extensions of SSIM (2)



- **Complex wavelet SSIM**
  - Motivation: robust to translation, rotation and scaling

$$\mathrm{SSIM}(\boldsymbol{x},\boldsymbol{y}) = \frac{2\left|\sum c_x \cdot c_y^*\right| + C}{\sum |c_x|^2 + \sum |c_y|^2 + C}$$

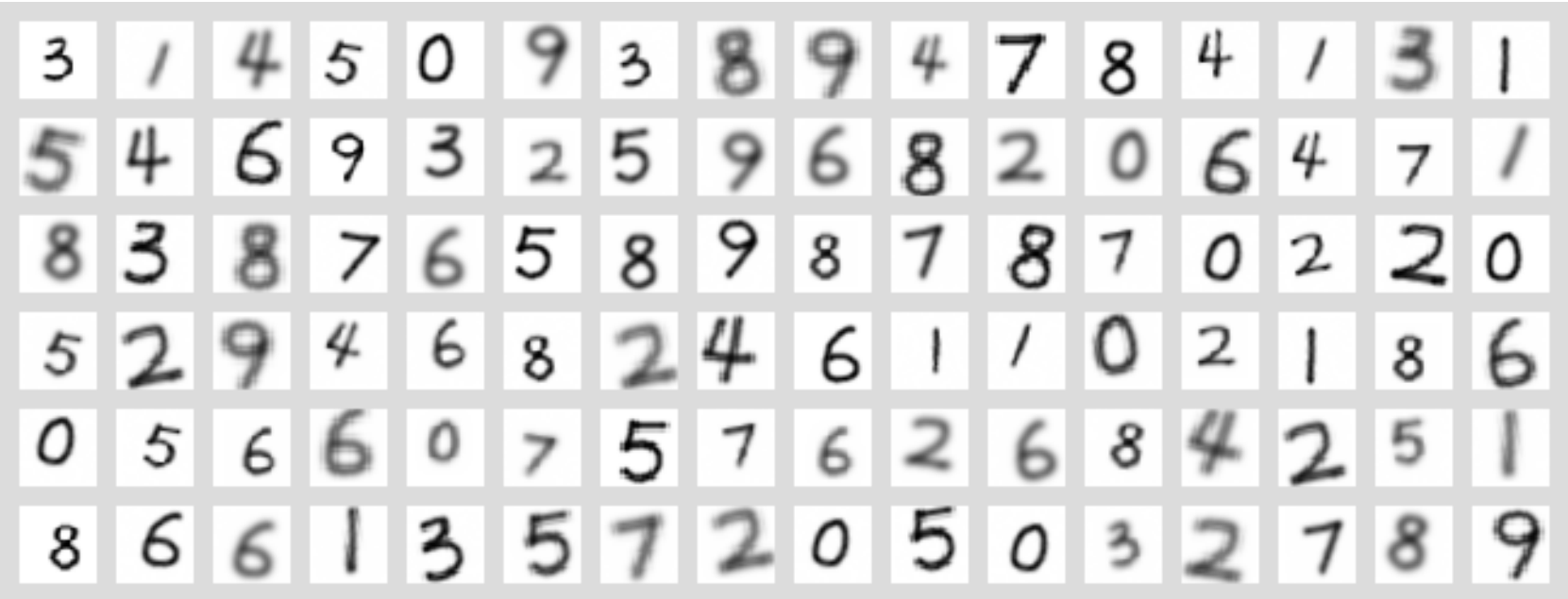
c_x, c_y : complex wavelet coefficients in images x and y

[Wang & Simoncelli, *ICASSP* '05]

Image Matching without Registration

Standard patterns: 10 images

Database: 2430 images



Correct Recognition Rate:

MSE: 59.6%; SSIM: 46.9%; Complex wavelet SSIM: **97.7%**

[Wang & Simoncelli, *ICASSP* '05]

Using SSIM

Web site: ***www.cns.nyu.edu/~lcv/ssim/***

SSIM Paper: **11,000+** downloads; Matlab code: **2400+** downloads

Industrial implementation: ***http://perso.wanadoo.fr/reservoir/***

- **Image/video coding and communications**
  - Image/video transmission, streaming & robustness [Kim & Kaveh '02, Halbach & Olsen '04, Lin *et al.* '04, Leontaris & Reibman '05]
  - Image/video compression [Blanch *et al.* '04, Dikici *et al.* '04 , Ho *et al.* '03, Militzer *et al.* '03]
  - High dynamic range video coding [Mantiuk *et al.* '04]
  - Motion estimation/compensation [Monmarthe '04]
- **Biomedical image processing**
  - Microarray image processing for bioinformatics [Wang *et al.* '03]
  - Image fusion of CT and MRI images [Piella & Heijmans '03, Piella '04]
  - Molecular image processing [Ling *et al.* '02]
  - Medical image quality analysis [Chen *et al.* '04]

Using SSIM (continued)

- **Watermarking/data hiding** [Alattar '03, Noore *et al.* '04, Macq *et al.* '04 Zhang & Wang '05, Kumsawat *et al.* '04]
- **Image denoising** [Park & Lee '04, Yang & Fox '04 , Huang *et al.* '05 Roth & Black '05, Hirakawa & Parks '05]
- **Image enhancement** [Battiato *et al.* '03]
- **Image/video hashing** [Coskun & Sankur '04, Hsu & Lu '04]
- **Image rendering** [Bornik *et al.* '03]
- **Image fusion** [Zheng *et al.* '04, Tsai '04, Gonzalez-Audicana *et al.* '05]
- **Texture reconstruction** [Toth '04]
- **Image halftoning** [Evans & Monga '03, Neelamani '03]
- **Radar imaging** [Bentabet '03]
- **Infrared imaging** [Torres '03, Pezoa *et al.* '04]
- **Ultrasound imaging** [Loizou *et al.* '04]
- **Vision processor design** [Cembrano *et al.*, '04]
- **Wearable display design** [von Waldkirch *et al.* '04]
- **Contrast equalization for LCD** [Iranli *et al.* '05]
- **Airborne hyperspectral imaging** [Christophe *et al.* '05]
- **Superresolution for remote sensing** [Rubert *et al.* '05]

THE END

Thank you!