EE 604
Digital Image Processing

IIT KANPUR
Indian Institute of Technology, Kanpur

Tanaya Guha
Aug - Nov 2017

Image Quality Assessment

Image quality

- **Image fidelity:** measure of (dis)similarity between two images or amount of error/distortion
- **Image quality:** measure of preference of one image over another
- If one of the images is a clean original, and the other is distorted, then **fidelity = quality.**
- Most popular image fidelity (quality) measure: **Mean Squared Error (MSE)**

MSE

$$\mathbf{x} = \{x_i | i = 1, 2, ...N\} \qquad \mathbf{y} = \{y_i | i = 1, 2, ...N\}$$

$$\mathrm{MSE}(\mathbf{x}, \mathbf{y}) = \frac{1}{N} \sum_{i=1}^{N} (x_i - y_i)^2$$

- A more generic form of MSE

$$d_p(\mathbf{x}, \mathbf{y}) = \left(\sum_{i=1}^{N} |e_i|^p \right)^{1/p} \quad \text{where} \quad e_i = x_i - y_i$$

- PSNR

$$\mathrm{PSNR} = 10 \log_{10} \frac{L^2}{\mathrm{MSE}} \quad \text{where L is the dynamic range}$$

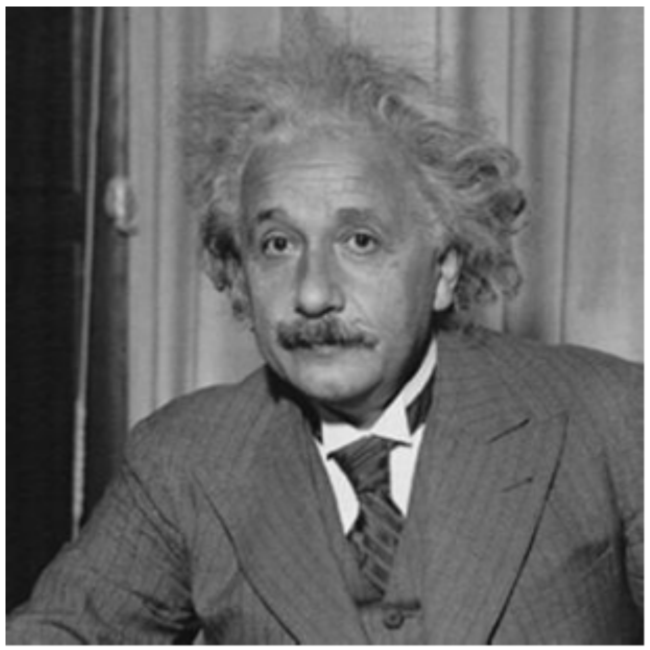
Why is MSE popular?

- Simple, parameter-free, non-expensive, memoryless
- Norm-based distance metric

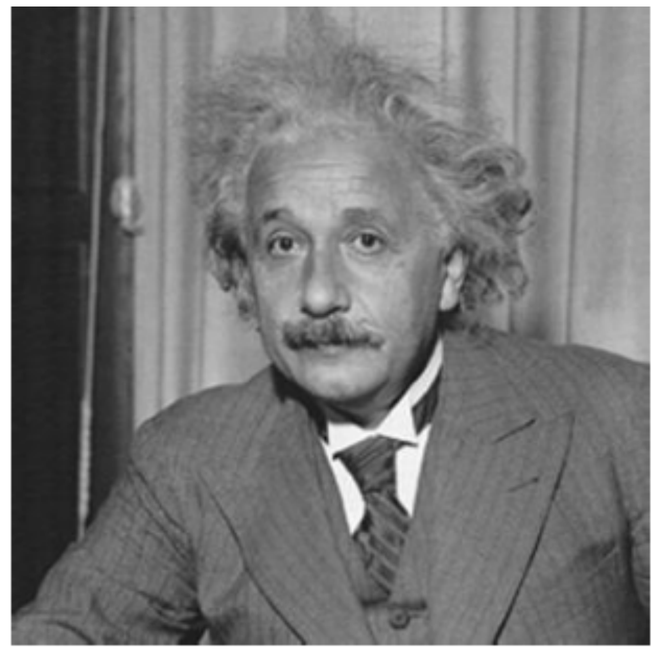
  — nonnegativity: $d_p(\mathbf{x}, \mathbf{y}) \geq 0$
  — identity: $d_p(\mathbf{x}, \mathbf{y}) = 0$ if and only if $\mathbf{x} = \mathbf{y}$
  — symmetry: $d_p(\mathbf{x}, \mathbf{y}) = d_p(\mathbf{y}, \mathbf{x})$
  — triangular inequality: $d_p(\mathbf{x}, \mathbf{z}) \leq d_p(\mathbf{x}, \mathbf{y}) + d_p(\mathbf{y}, \mathbf{z})$

- Natural relation to energy
- Excellent for optimization: convex, symmetric, differentiable

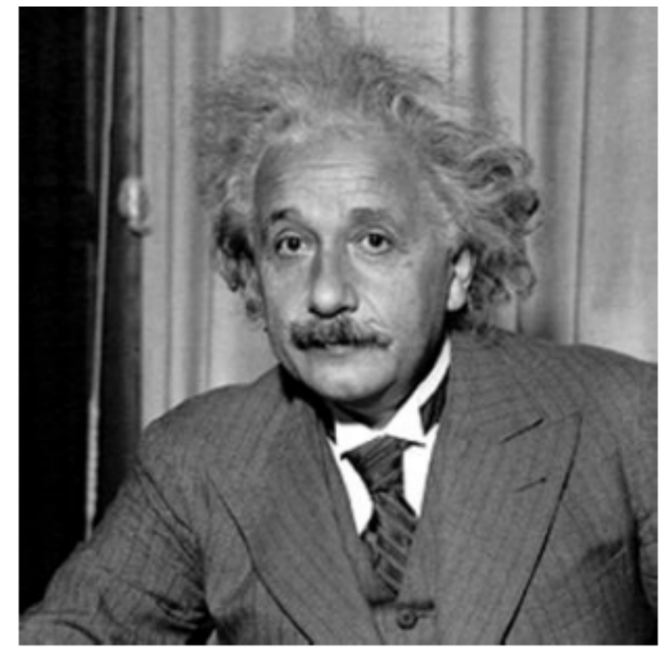
What's wrong with MSE?



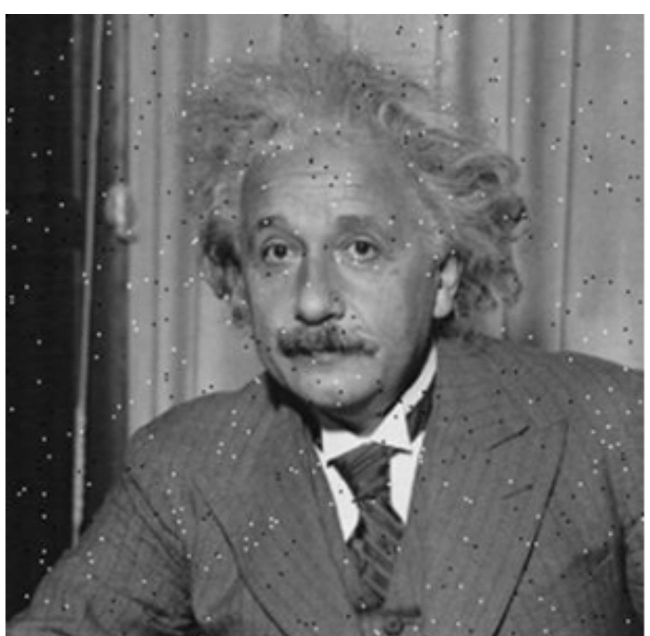
Original



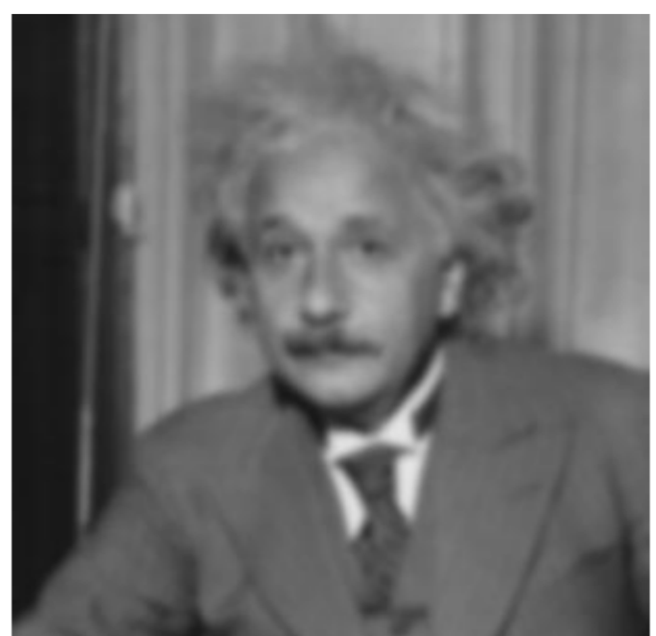
MSE = 144



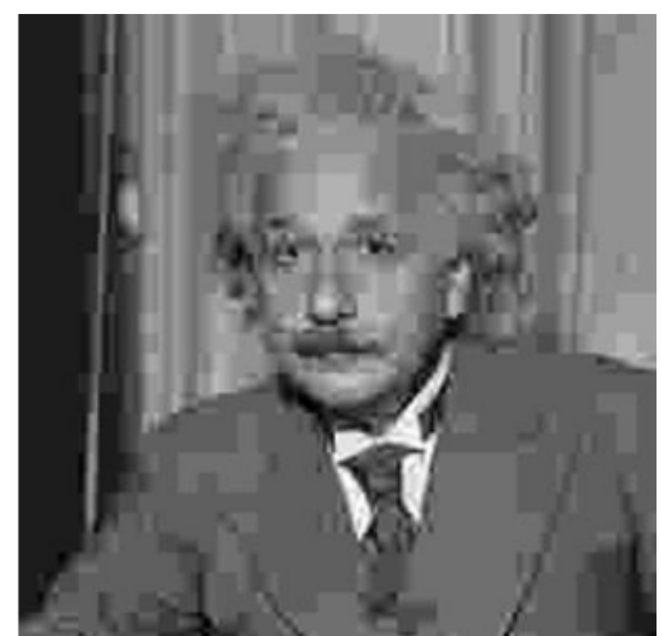
MSE = 144



MSE = 144



MSE = 144

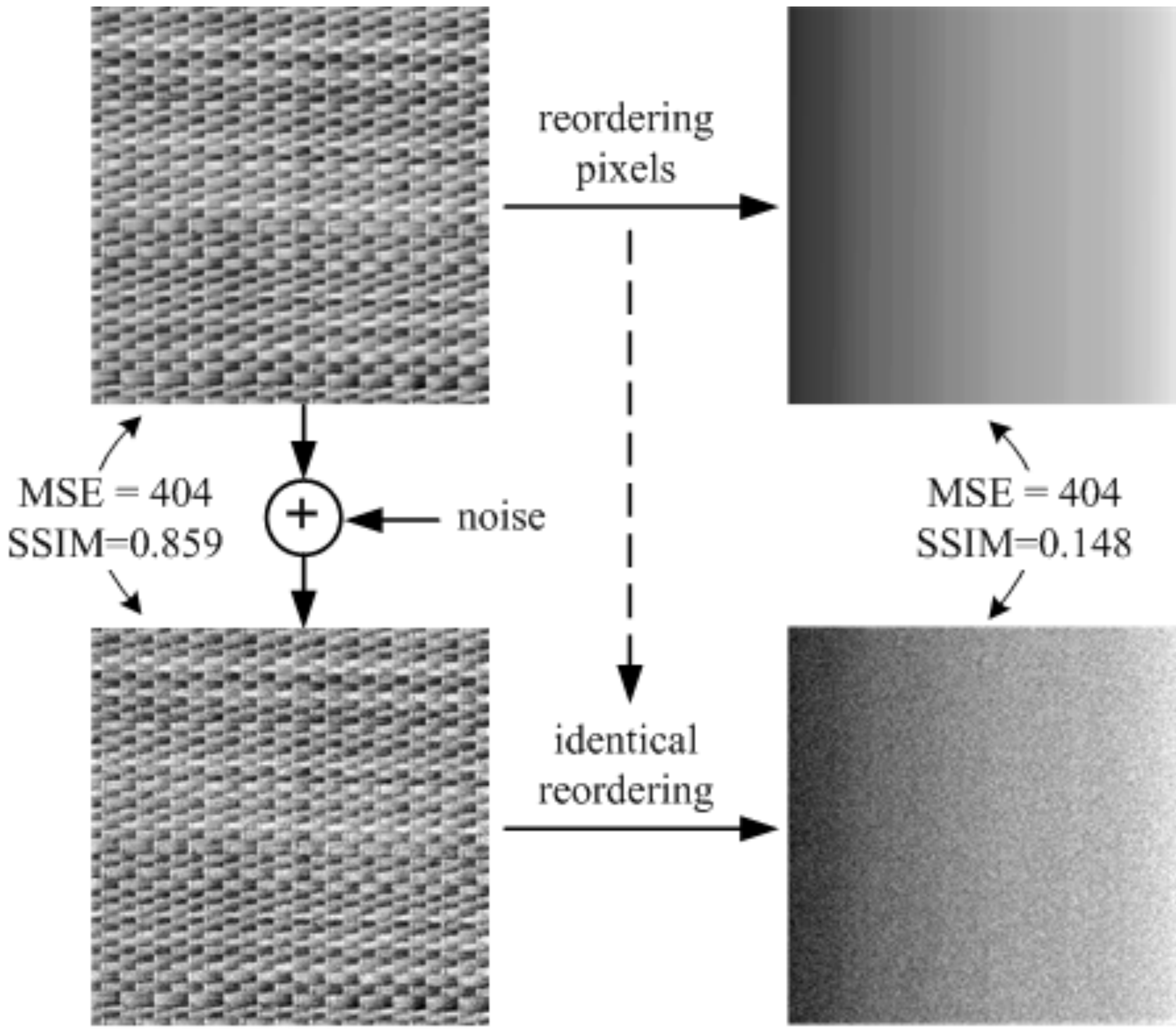


MSE = 142

Why does MSE fail?

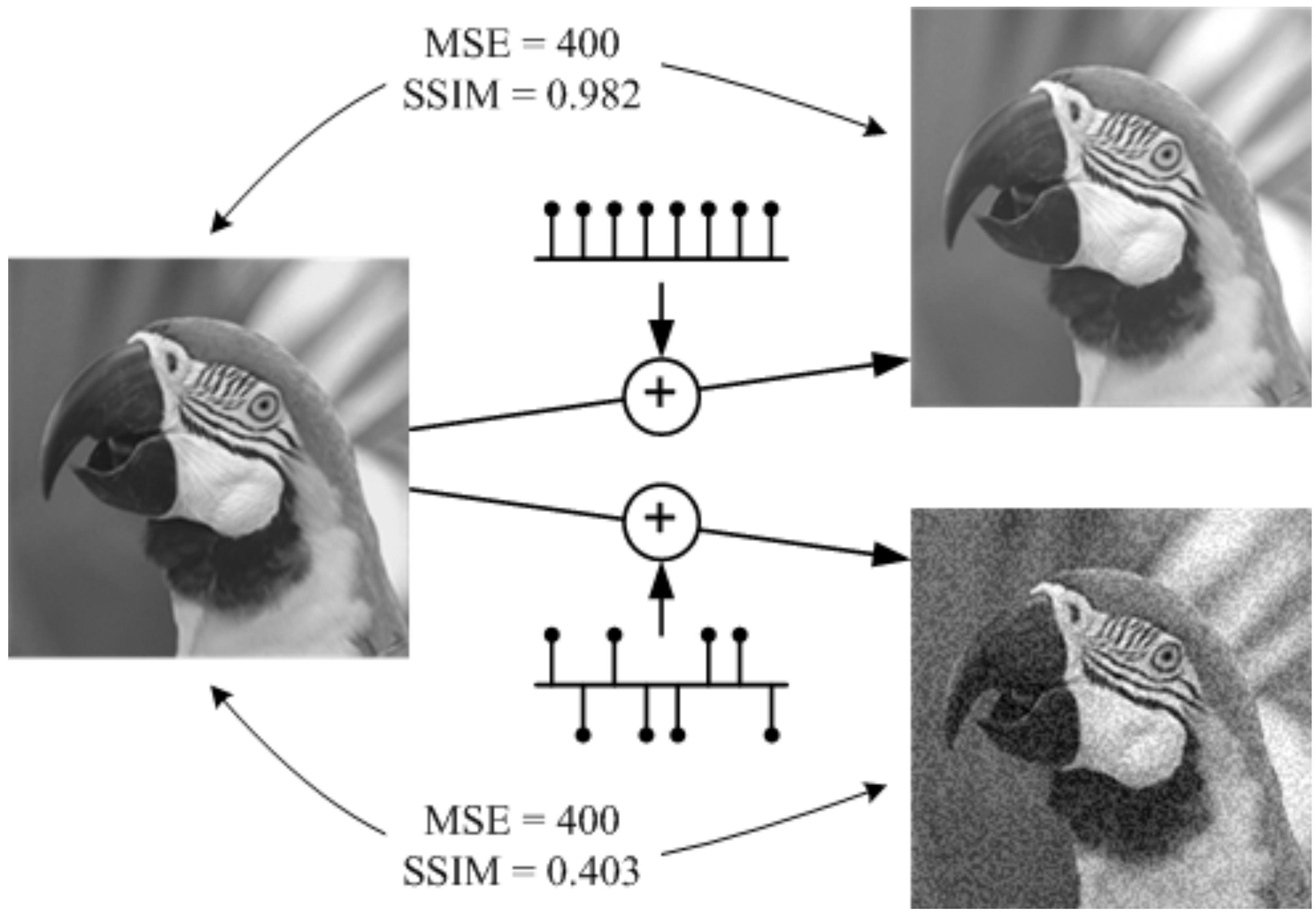
- Several strong assumptions
  - Fidelity measure is spatially independent
  - All points are equally important for fidelity
  - Sign of error does not matter
  - The error and the original image has no relationship.

Why does MSE fail?



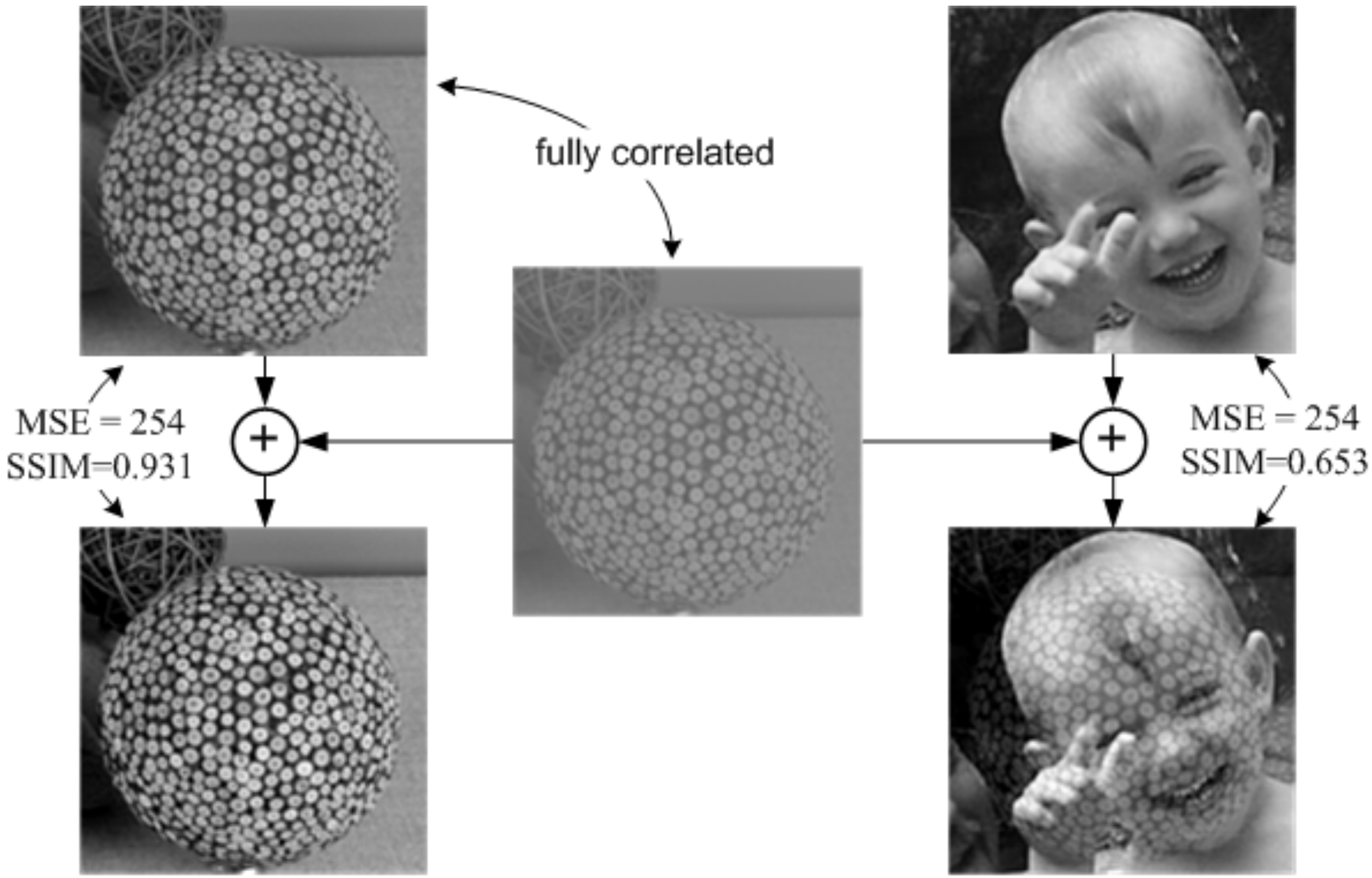
[Wang and Bovik 2009]

Why does MSE fail?



[Wang and Bovik 2009]

Why does MSE fail?



[Wang and Bovik 2009]

Perceptual Image quality

- Image processing systems need to measure how good the output image is.
- The end users of many image processing systems are often humans, so human perception of quality is an important criteria.
- MSE does not correlate well with visual perception of quality.
- **Solution?**
    - Try to model HVS!
        - Difficult due to the complexity and our relatively less understanding of HVS
    - Develop metrics **based on HVS properties**
        - These metrics correlate better with human perception. These metrics are often called **perceptual image quality metrics**.

SSIM Index

Image quality assessment: from error visibility to structural similarity

View Document

4 Author(s)

Zhou Wang ; A.C. Bovik ; H.R. Sheikh ; E.P. Simoncelli

View All Authors

Abstract | Authors | Figures | References | Citations | Keywords | Metrics | Media

Abstract:
Objective methods for assessing perceptual image quality traditionally attempted to quantify the visibility of errors (differences) between a distorted image and a reference image using a variety of known properties of the human visual system. Under the assumption that human visual perception is highly adapted for extracting structural information from a scene, we introduce an alternative complementary framework for quality assessment based on the degradation of structural information. As a specific example of this concept, we develop a structural similarity index and demonstrate its promise through a set of intuitive examples, as well as comparison to both subjective ratings and state-of-the-art objective methods on a database of images compressed with JPEG and JPEG2000. A MATLAB implementation of the proposed algorithm is available online at http://www.cns.nyu.edu//spl sim/lcv/ssim/.

Published in: IEEE Transactions on Image Processing (Volume: 13, Issue: 4, April 2004)

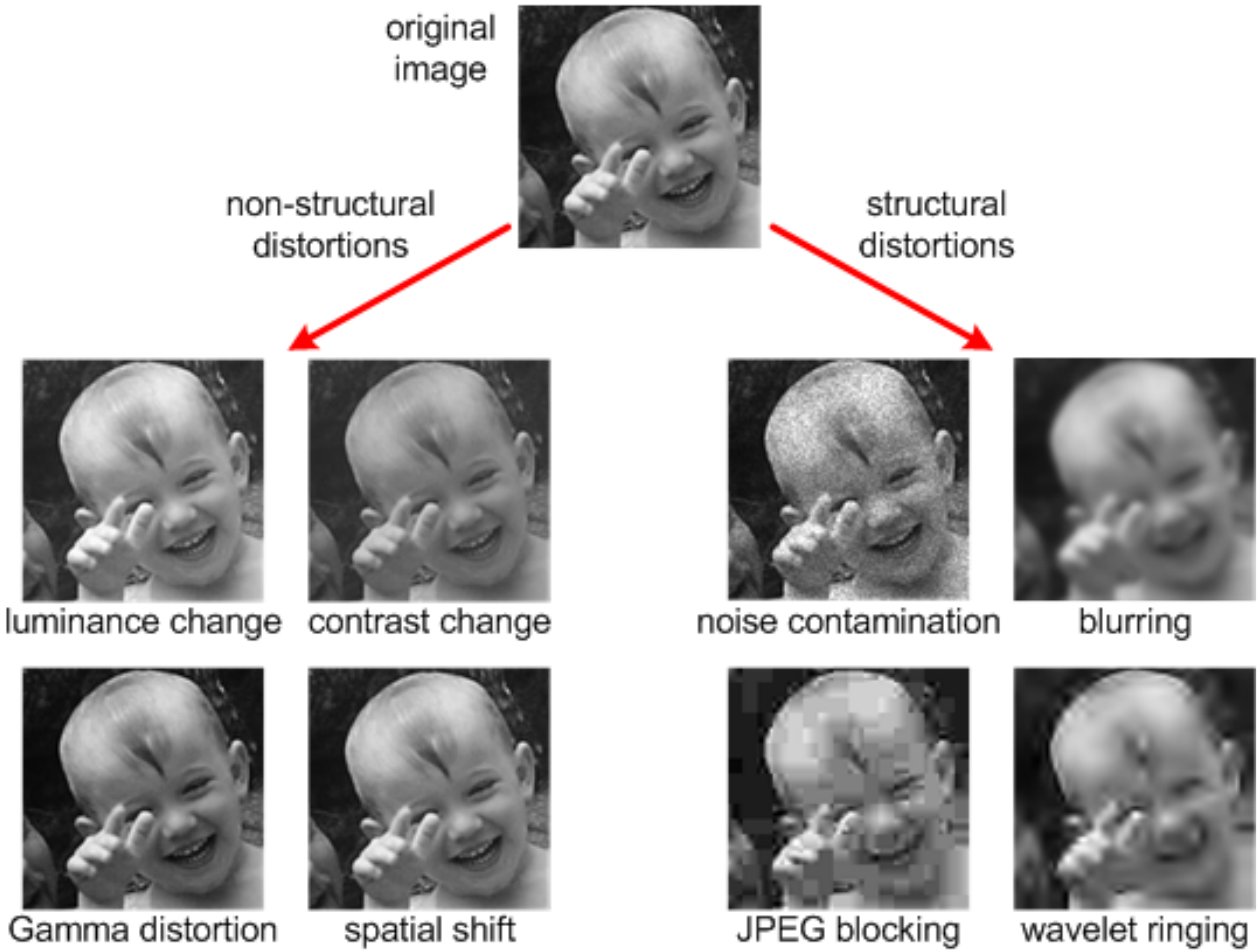
SSIM Index [Wang et al. 2004]

- **Generic:** does not assume any prior knowledge of distortion type
- **Full reference:** assumes the availability of a full information of the image being compared with (reference image)
- **opinion-unaware:** does not use human scores for computing quality
- **Assumption:** HVS has evolved to and extract information from structures. HVS is more sensitive to structural-distortions than non-structural distortions.
- SSIM separates structural and non-structural distortions

Structural vs. Non-structural

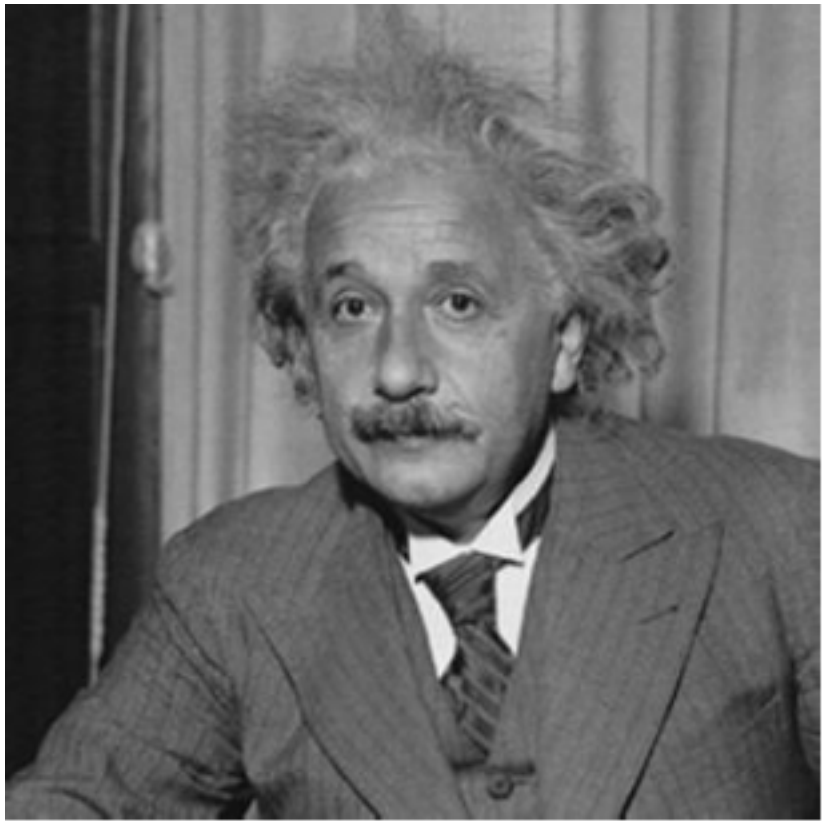
- What do we understand by image **“structures”**?
    - spatial patterns, repetitive patterns
    - edges, shapes, corners
- **Structural distortion**
    - results in changes in image structure i.e. loss/addition/modification of structural information
        - caused by blurring, noise, motion, compression etc.
- **Non-structural distortion**
    - results in changes in illumination, contrast, color or shift

Structural vs. Non-structural

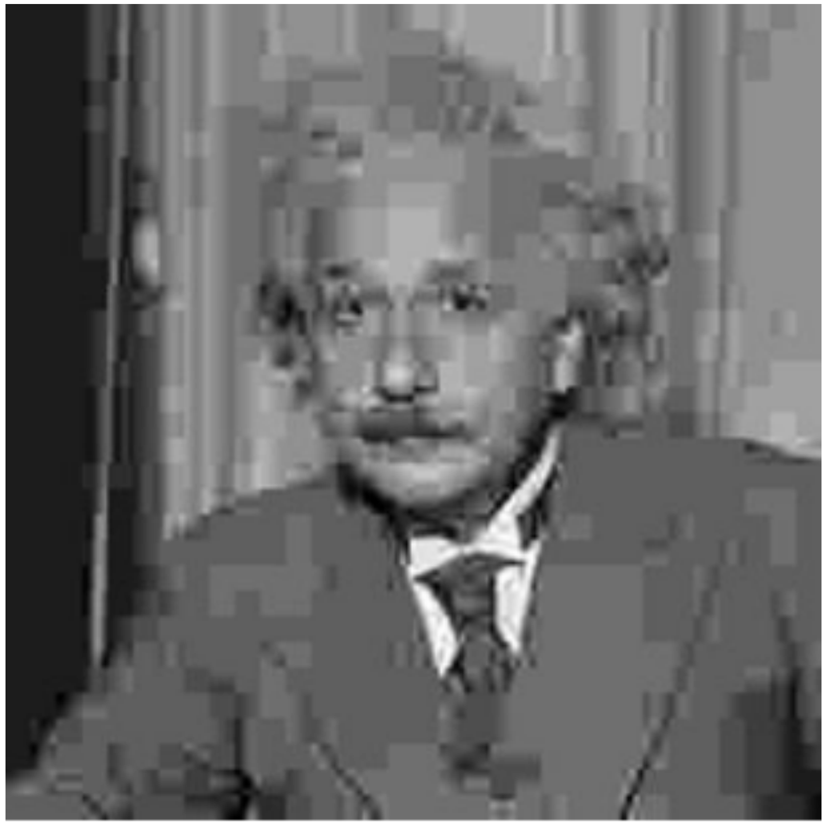


[Wang and Bovik 2009]

SSIM Index



Reference image: **X**



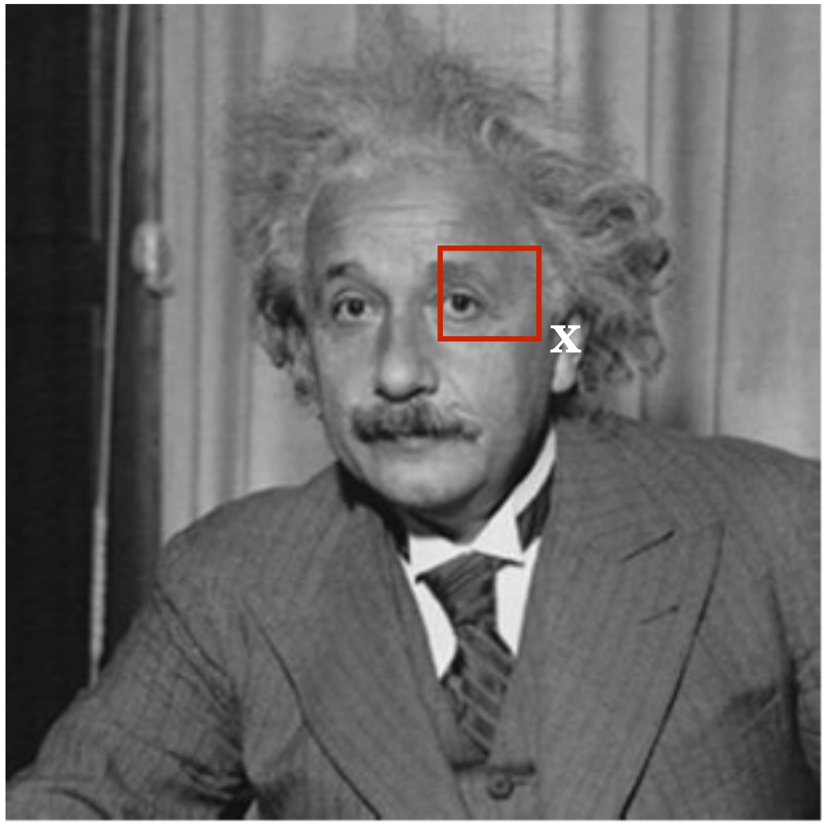
Distorted image: **Y**

Compute quality of **Y** with respect to **X**

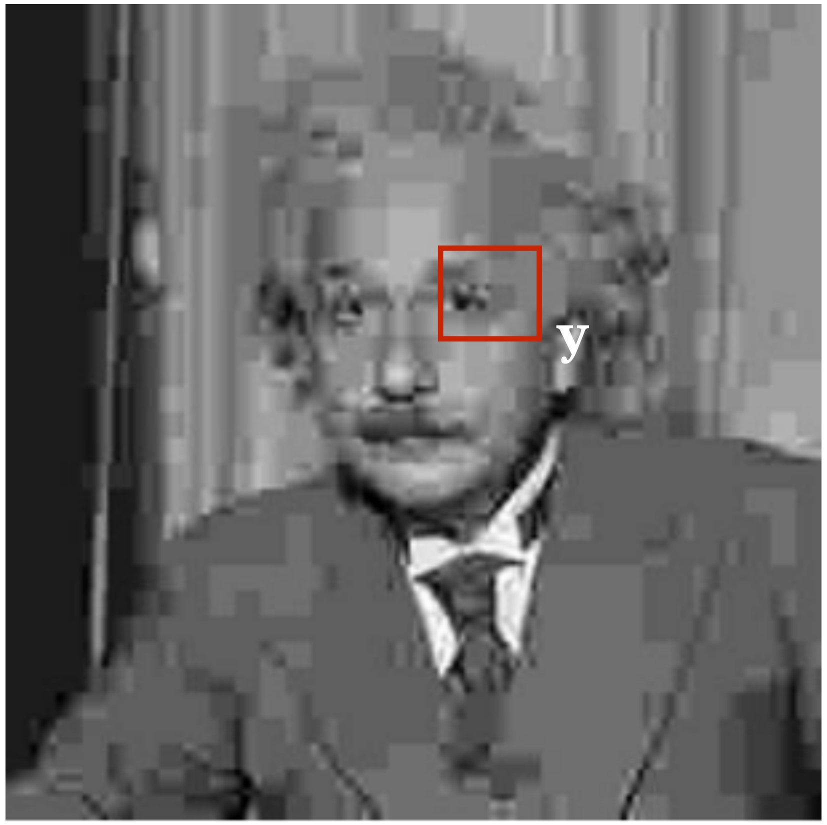
SSIM Index

- SSIM compares two images at pixel level
    - A square patch centered at pixel $\mathbf{x}$ in $\mathbf{X}$ is compared with a square patch $\mathbf{y}$ in $\mathbf{Y}$

SSIM Index



Reference image: **X**



Distorted image: **Y**

Compute quality of **Y** with respect to **X**

SSIM Index

Luminance comparison:

- local luminance is modeled by **mean intensity** of the local region

$\mathbf{x}$

$\downarrow$

μ_{x}

$\mathbf{y}$

$\downarrow$

μ_{y}

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

SSIM Index

Contrast comparison:

- local contrast is modeled by **standard deviation** of intensity of the local region

$\mathbf{x}$

σ_{x}

$\mathbf{y}$

σ_{y}

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

SSIM Index

Structural comparison:

- Remove the non-structural distortion by normalization
- Structural similarity modeled by correlation between normalized patches

x

$$\sigma_{xy} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu_x)(y_i - \mu_y)$$

y

$$(\mathrm{x} - \mu_{\mathrm{x}})/\sigma_{\mathrm{x}} \longleftarrow \text{cosine similarity} \longrightarrow (\mathrm{y} - \mu_{\mathrm{y}})/\sigma_{\mathrm{y}}$$

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

SSIM Index

x 

 **y**

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

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

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

$$\mathrm{SSIM}(\mathbf{x},\mathbf{y}) = [l(\mathbf{x},\mathbf{y})]^\alpha \cdot [c(\mathbf{x},\mathbf{y})]^\beta \cdot [s(\mathbf{x},\mathbf{y})]^\gamma$$

where $\alpha, \beta, \gamma > 0$

SSIM Index

- **Local SSIM score**

$$\mathrm{SSIM}(\mathbf{x},\mathbf{y}) = [l(\mathbf{x},\mathbf{y})]^{\alpha} \cdot [c(\mathbf{x},\mathbf{y})]^{\beta} \cdot [s(\mathbf{x},\mathbf{y})]^{\gamma}$$

x **y**

- **For** $\alpha = \beta = \gamma = 1 \qquad C_3 = C_2/2$

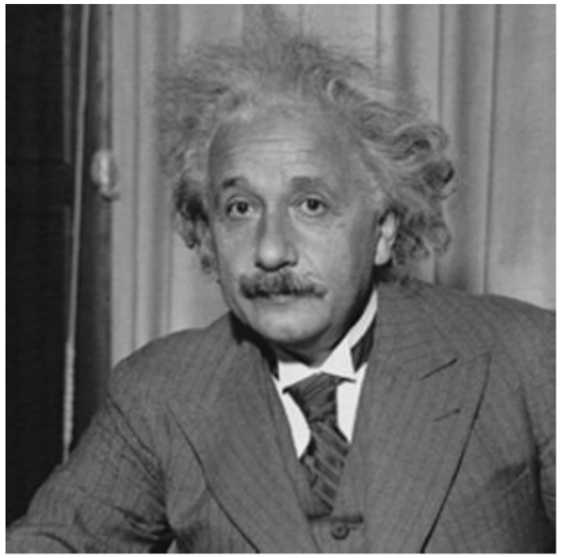
$$\mathrm{SSIM}(\mathbf{x},\mathbf{y}) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{\left(\mu_x^2 + \mu_y^2 + C_1\right)\left(\sigma_x^2 + \sigma_y^2 + C_2\right)}.$$

SSIM Index

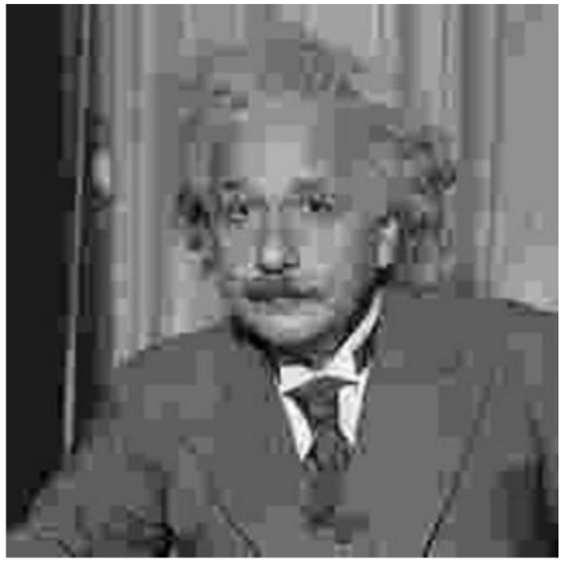
Properties:

- <u>Symmetry:</u> SSIM(**x**,**y**) = SSIM(**y**,**x**)
- <u>Boundedness:</u> SSIM(**x**,**y**)<= 1
- <u>Unique maximum:</u> SSIM(**x**,**y**) = 1 only when **x=y**

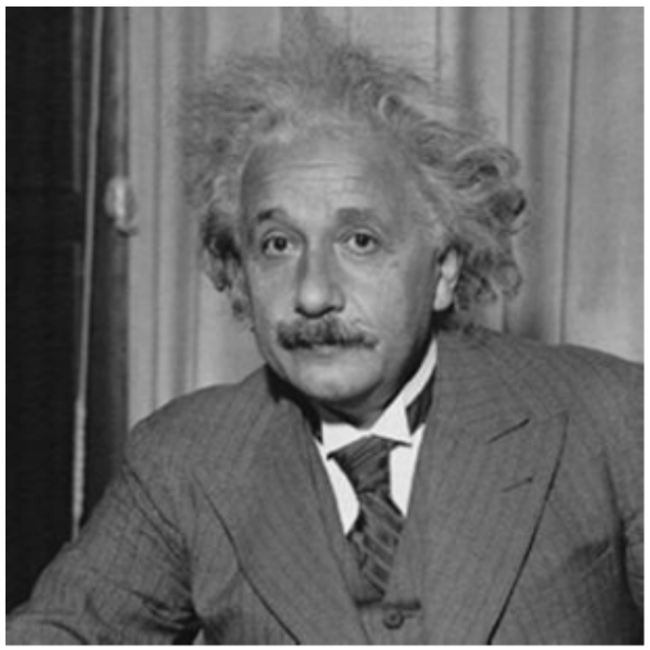
Global SSIM score



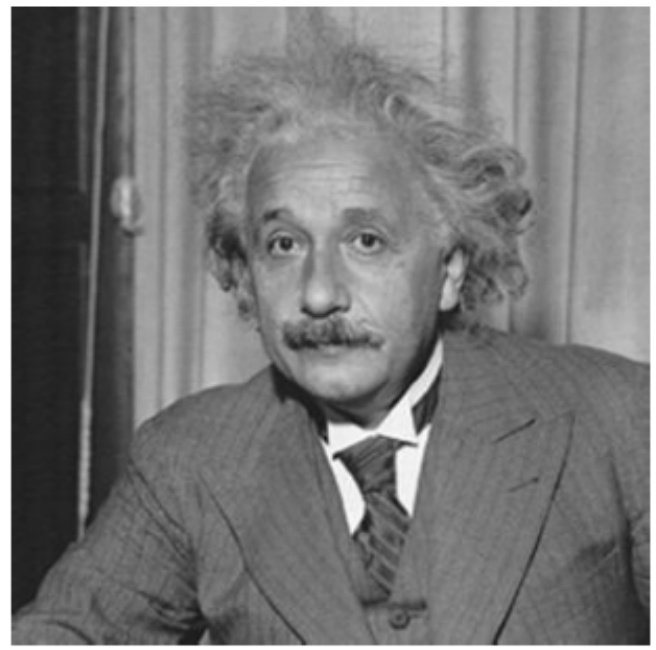
$$\mathrm{MSSIM}(\mathbf{X}, \mathbf{Y}) = \frac{1}{M} \sum_{j=1}^{M} \mathrm{SSIM}(\mathbf{x}_j, \mathbf{y}_j)$$



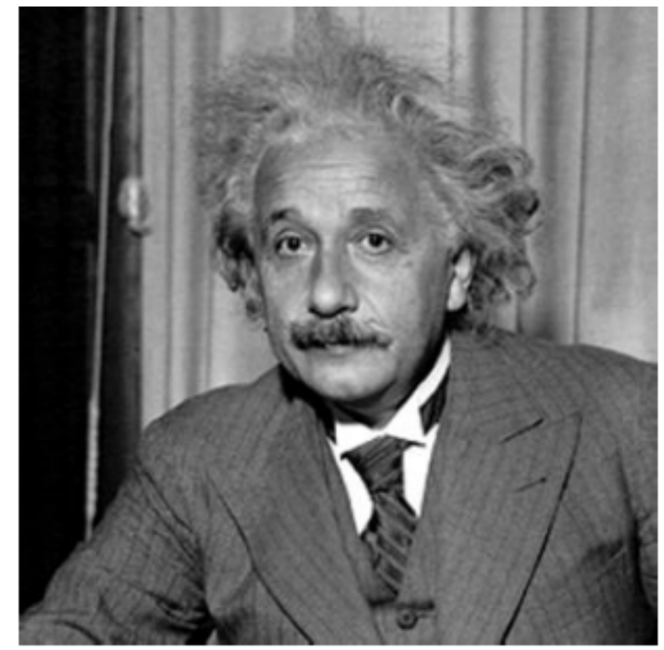
Results



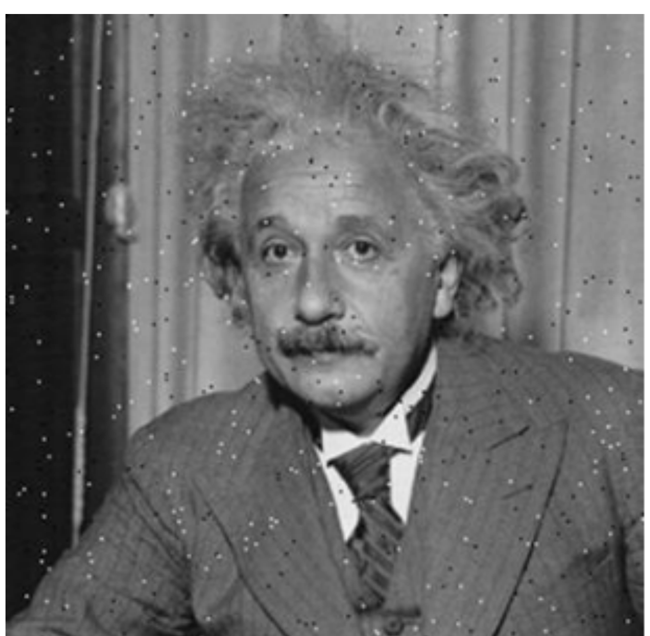
Original



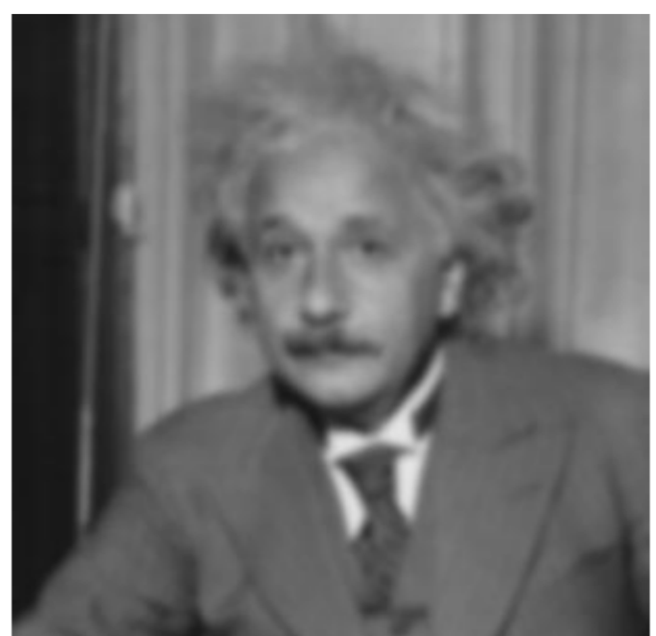
SSIM = 1



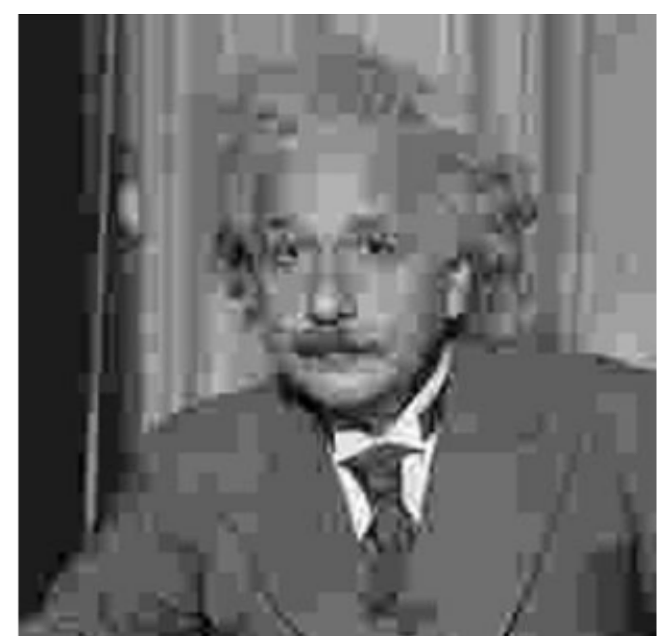
SSIM=0.988



SSIM=0.913



SSIM=0.694



SSIM=0.692

Pooling strategies in SSIM

- The easiest is **mean pooling** (vanilla SSIM)
- **Minkowski pooling**

$$s_i = \mathrm{SSIM}(\mathbf{x}_i, \mathbf{y}_i) \qquad \mathrm{S}(\mathbf{X}, \mathbf{Y}) = \frac{1}{M}\sum_{i=1}^{N}(s_i)^p$$

Pooling strategies in SSIM

- The easiest is **mean pooling** (vanilla SSIM)

- **Minkowski pooling**

$$s_i = \mathrm{SSIM}(\mathbf{x}_i, \mathbf{y}_i) \qquad \mathrm{S}(\mathbf{X}, \mathbf{Y}) = \frac{1}{M}\sum_{i=1}^{N}(s_i)^p$$

- **Distortion-weighted pooling**

$$\mathrm{S}(\mathbf{X}, \mathbf{Y}) = \frac{\sum_{i=1}^{M} w_i s_i}{\sum_{i=1}^{M} w_i} \qquad w_i = f(s_i)$$

$$\text{e.g. } w_i = |s_i|^4$$

  - Areas of higher distortion contributes more to visual quality

Pooling strategies in SSIM

- **Information content-weighted pooling**

$$\mathrm{S}(\mathbf{X}, \mathbf{Y}) = \frac{\sum_{i=1}^{M} w_i s_i}{\sum_{i=1}^{M} w_i} \qquad w_i = g(\mathbf{x}_i, \mathbf{y}_i)$$

- Result from information theory: $I = \frac{1}{2}\log(1 + \frac{S}{C})$
    - I = information content, S = signal power, C = noise power

$$g(\mathbf{x}, \mathbf{y}) = \log[(1 + \frac{\sigma_x^2}{C})(1 + \frac{\sigma_y^2}{C})]$$

[Wang & Shang 2006]