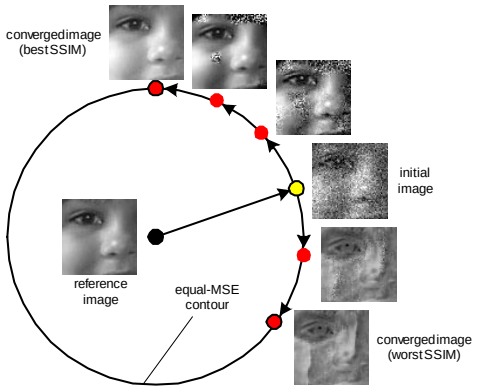
Multimedia Computing

Perceptual Image Quality Assessment

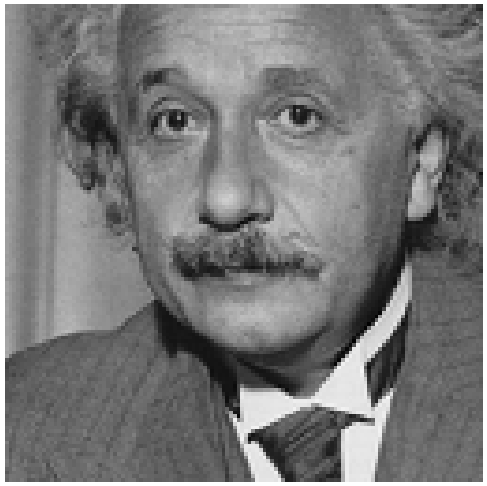


Outline

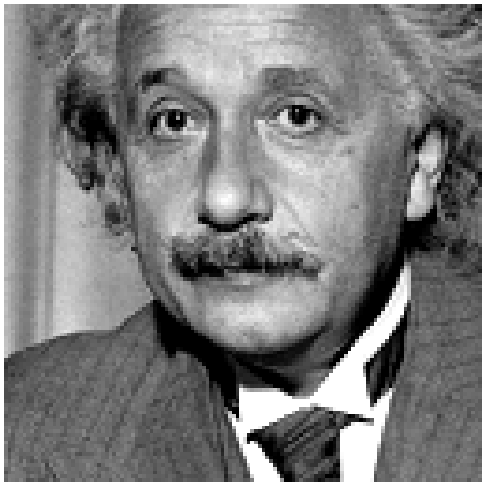
- Why “Perceptual”?
  - Motivation
  - Overview
- Perceptual Image Quality Assessment
  - Mean squared error’s deficiency
  - Error visibility methods
  - Structural similarity methods
- Perceptual Image Processing
  - Image compression
  - ……

Motivation

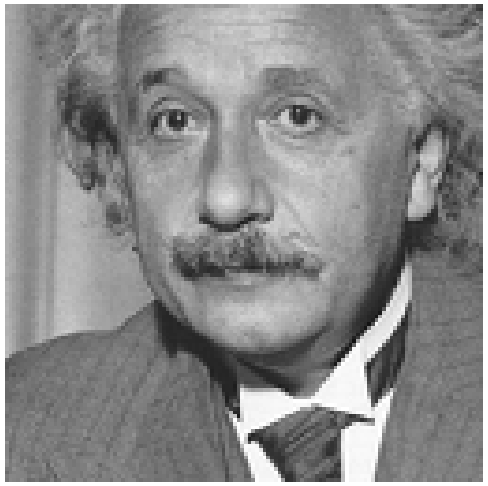
original Image



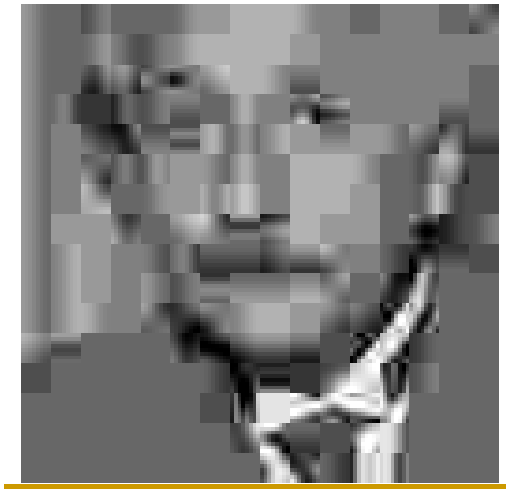
MSE=0, SSIM=1



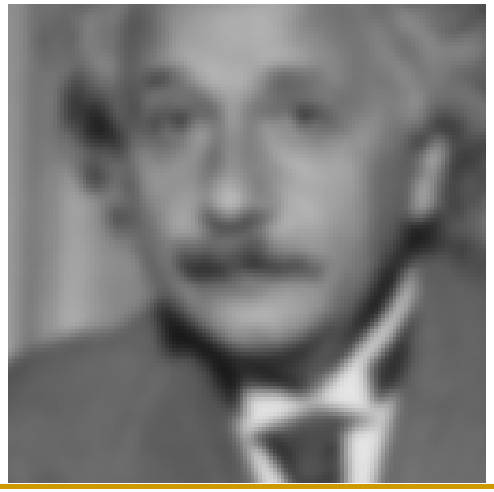
MSE=309, SSIM=0.928



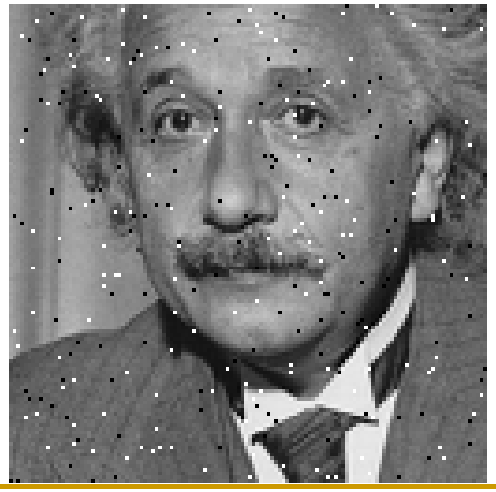
MSE=309, SSIM=0.987



MSE=309, SSIM=0.580



MSE=309, SSIM=0.641



MSE=309, SSIM=0.730

Perceptual Image Processing Overview

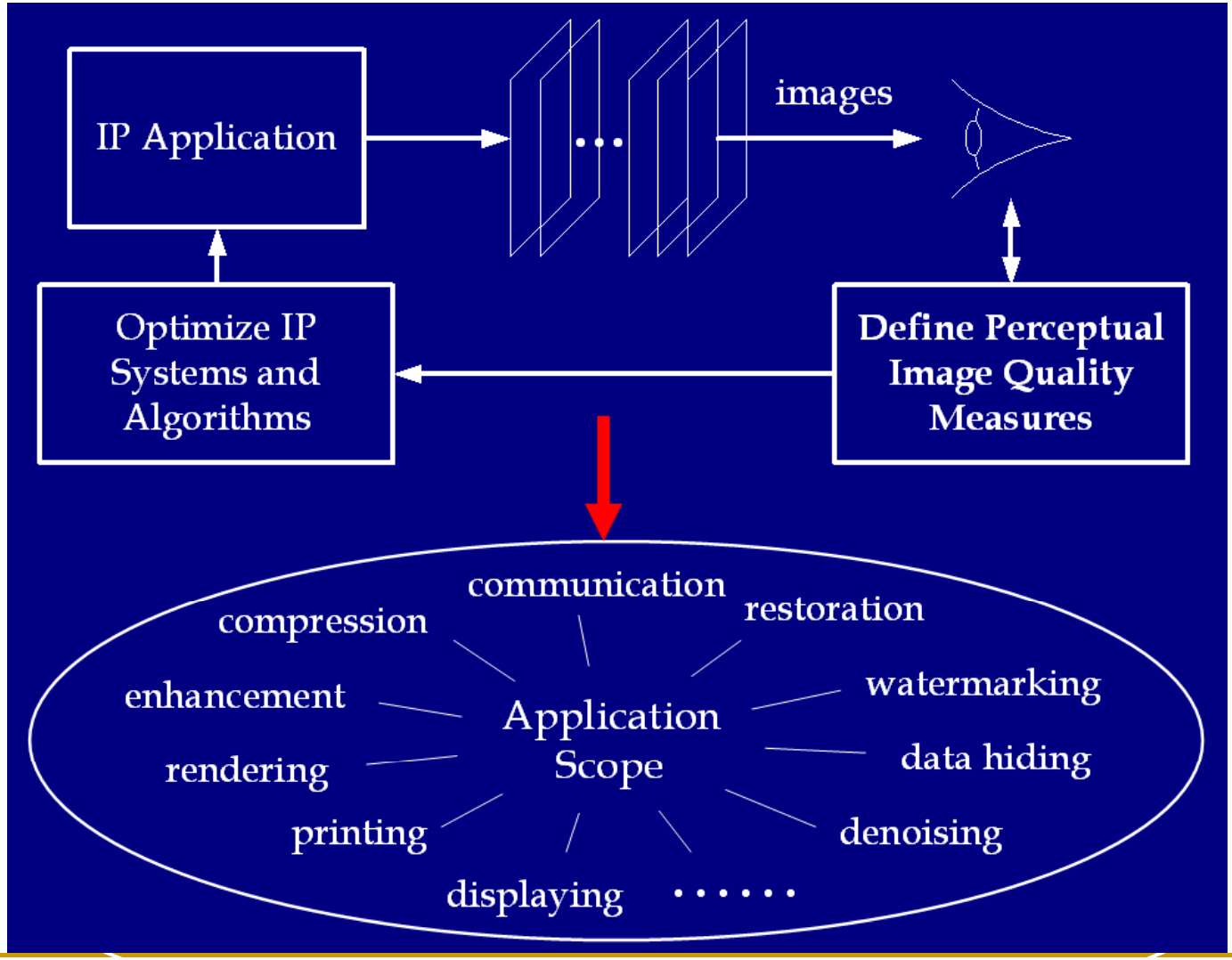


Image Quality Assessment: Classifications

	Full-reference	Reduced-reference	No-reference
Application-specific	Many	Some	Some
General-purpose	Many	Very Few	None

Widely used: MSE or PSNR

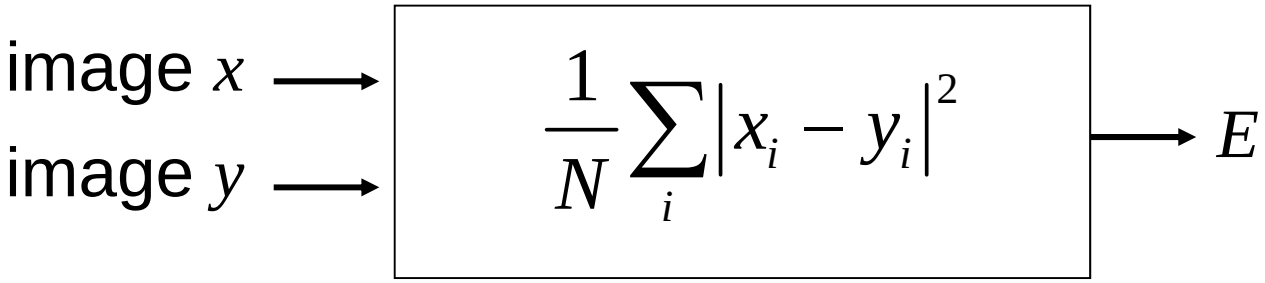
- Availability of Reference:
  - Full-Reference (FR): reference (original) image available
  - No-Reference (NR): reference image not available
  - Reduced-Reference (RR): reference image partially available
- Application Scope
  - General-purpose vs. application-specific

Outline

- Why “Perceptual”?
  - Motivation
  - Overview
- Perceptual Image Quality Assessment
  - Mean squared error’s deficiency
  - Error visibility methods
  - Structural similarity methods
- Perceptual Image Processing
  - Image compression
  - ……

Mean Squared Error's Deficiency (1)

- Mean Squared Error

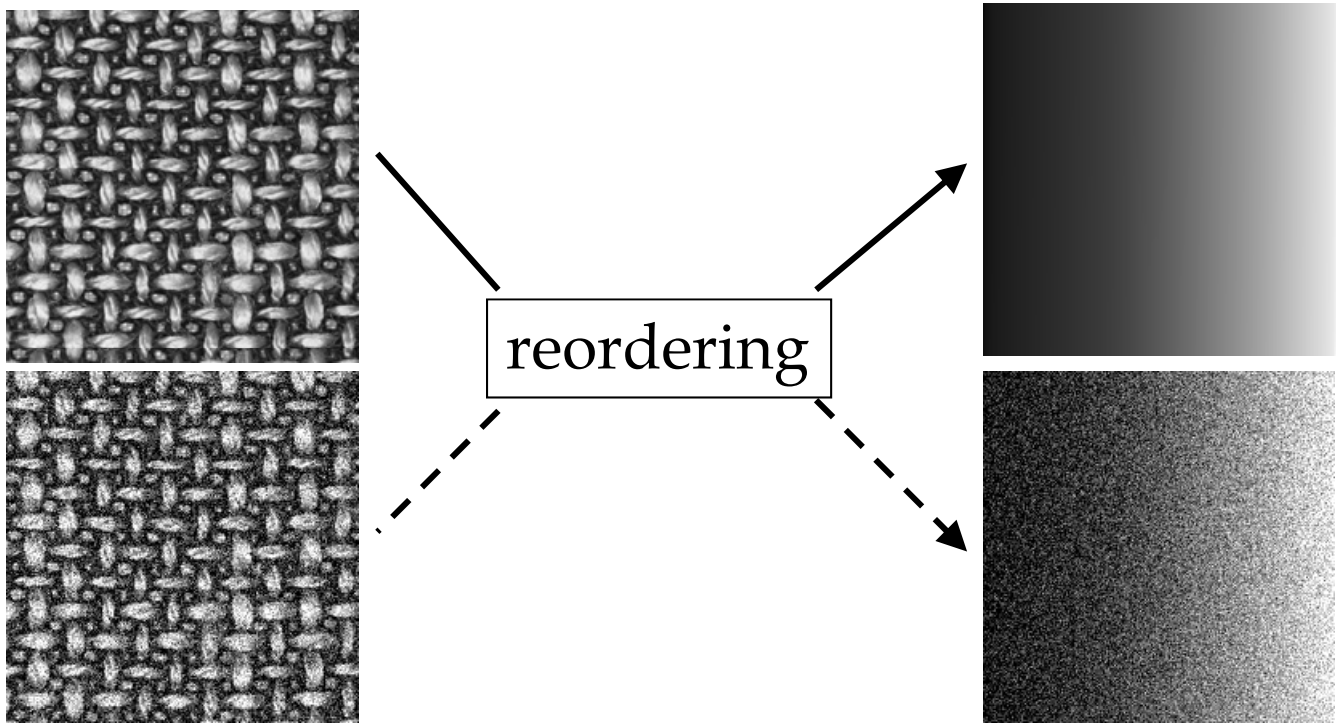


- Advantages
  - Easy to compute
  - Easy to optimize
  - Clear physical meaning: energy
- What's the problem?

Mean Squared Error's Deficiency (2)

$$\frac{1}{N}\sum_{i}\left|x_i - y_i\right|^2$$

Don't care about ordering



MSE = 1600, MSSIM = 0.6373

MSE = 1600, MSSIM = 0.0420

Mean Squared Error's Deficiency (3)

$$\frac{1}{N}\sum_i |x_i - y_i|^2$$

Don't care about the sign

+ 30

MSE = 900

SSIM = 0.9329

+ (rand sign)* 30

MSE = 900

SSIM = 0.2470

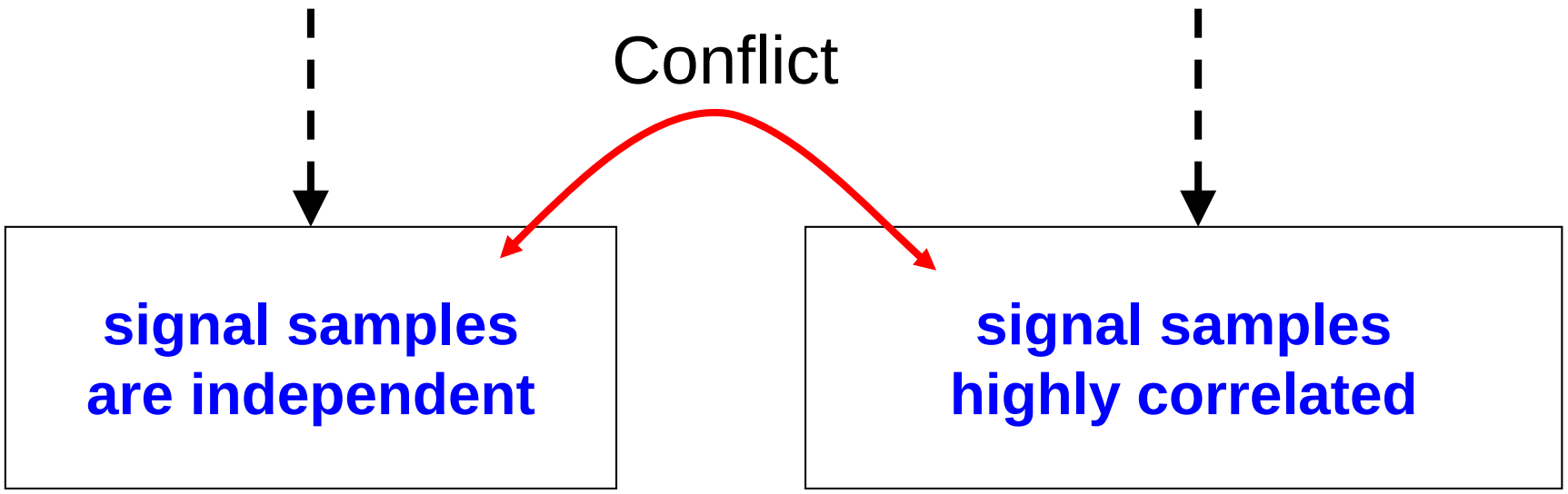
Mean Squared Error's Deficiency (4)

- Mean Squared Error

$$E = \frac{1}{N}\sum_i |x_i - y_i|^2$$

- Natural Images

highly structured



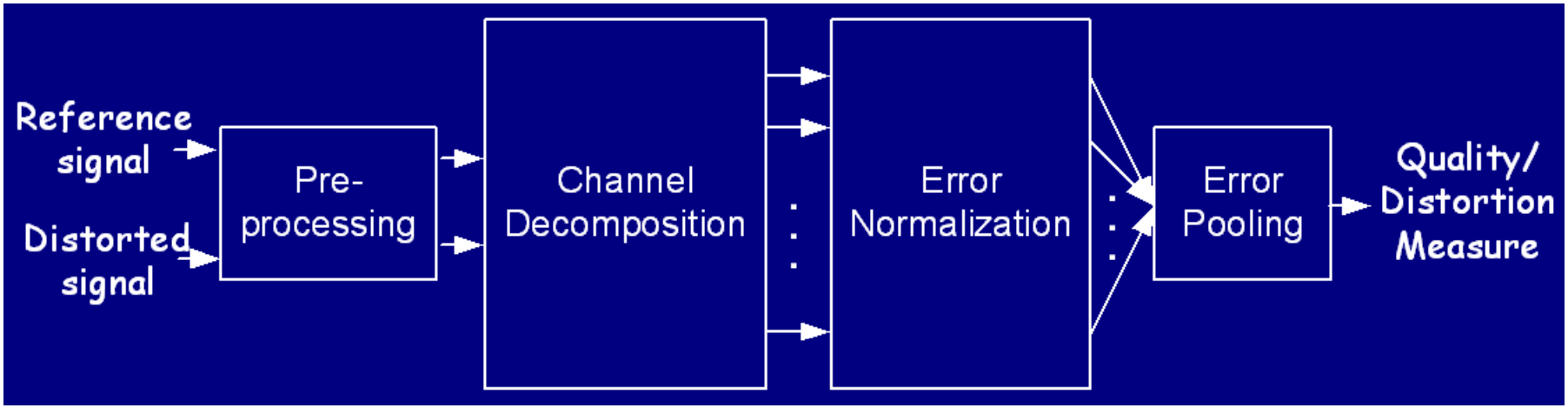
Error Visibility Method: Idea

distorted signal = reference signal + error signal

Quantify **error signal** **perceptually**

- **Representative work**
  - Frequency weighting (pioneering work) [Mannos & Sakrison, 74]
  - 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]
  - SSIM (Wang et al., 2004)
  - FSIM (Zhang et al., 2011)

Error Visibility Method: Framework



- Goal: simulate relevant early HVS components
  - Structures motivated by **physiology**
  - Parameters determined by **psychophysics**

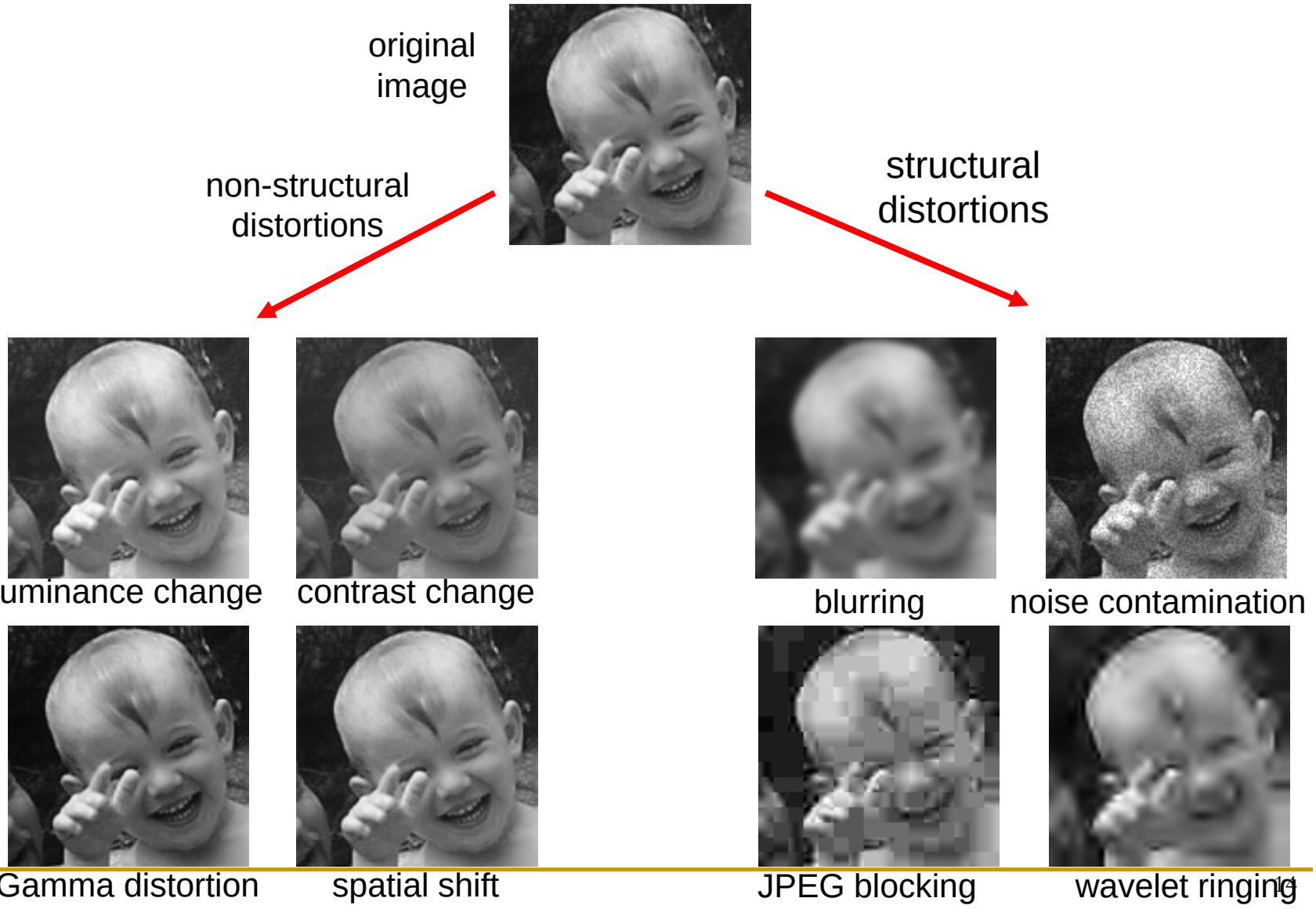
Structure Similarity Methods: Idea

Purpose of vision: extract structural information

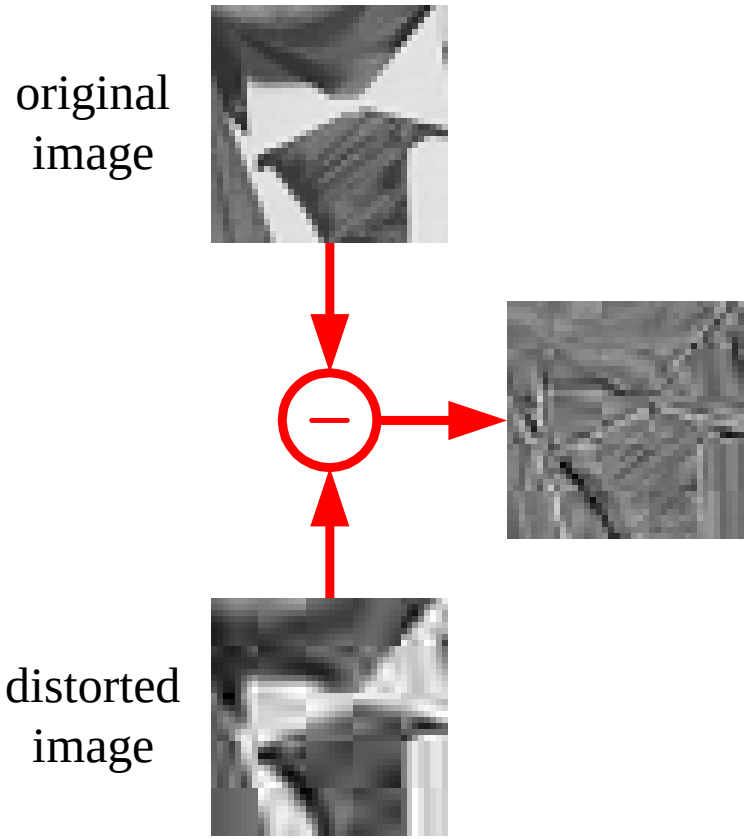
Quantify **structural distortion**

- **Questions:**
  - How to define structural/nonstructural distortions?
  - How to separate structural/nonstructural distortions?

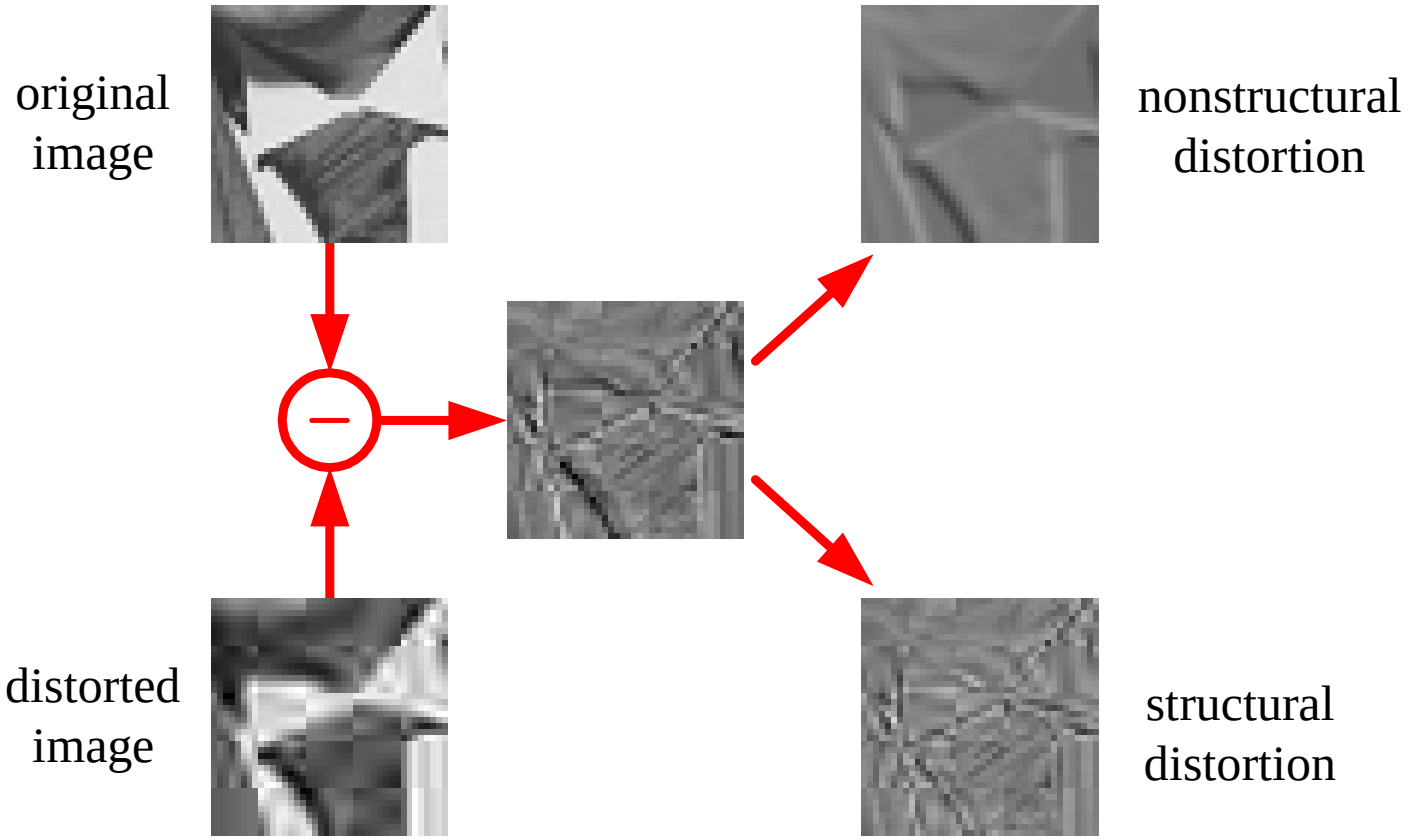
What are Structural/Non-Structural Distortions?



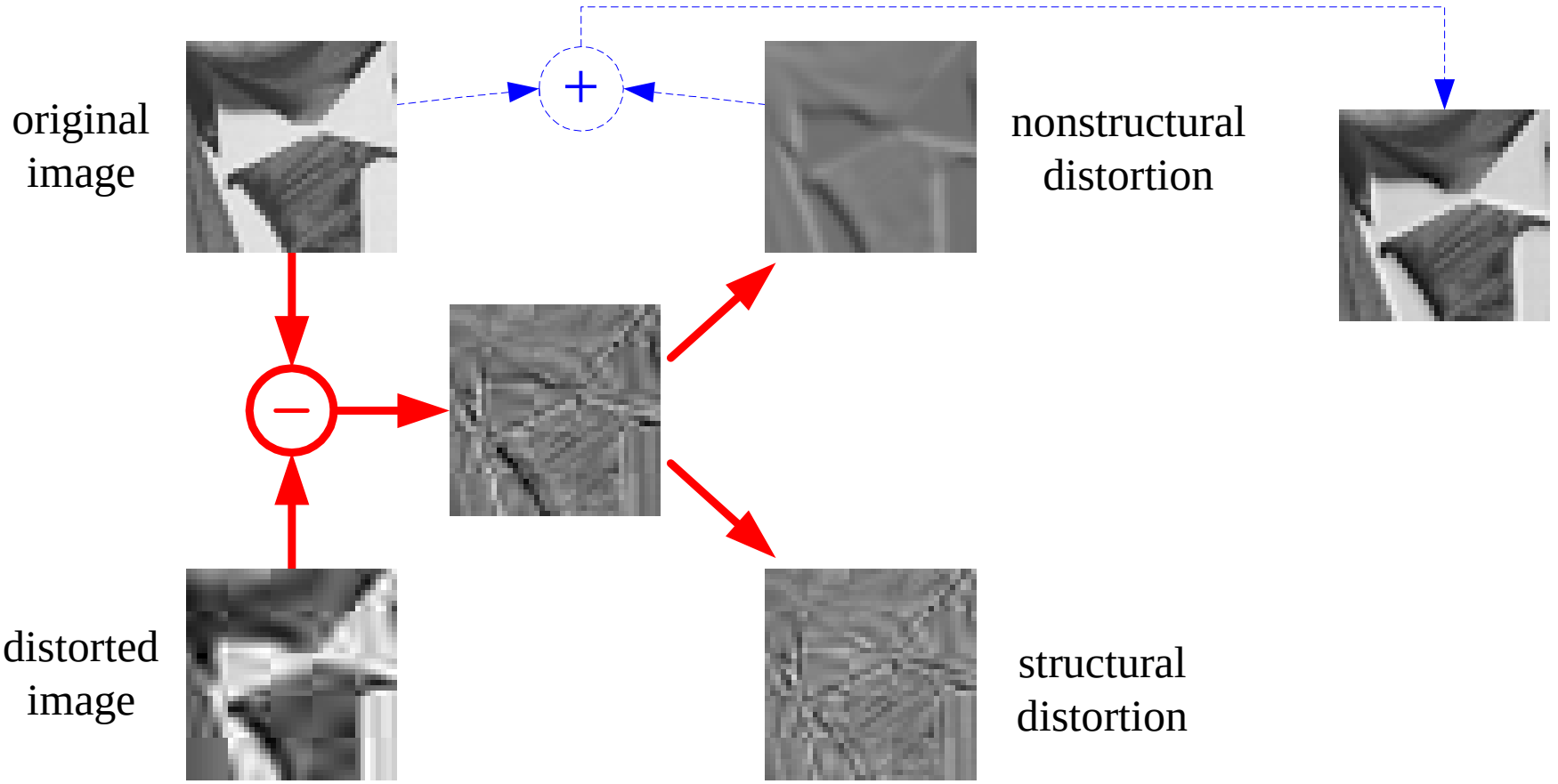
Separation of Structural/Non-Structural Distortions



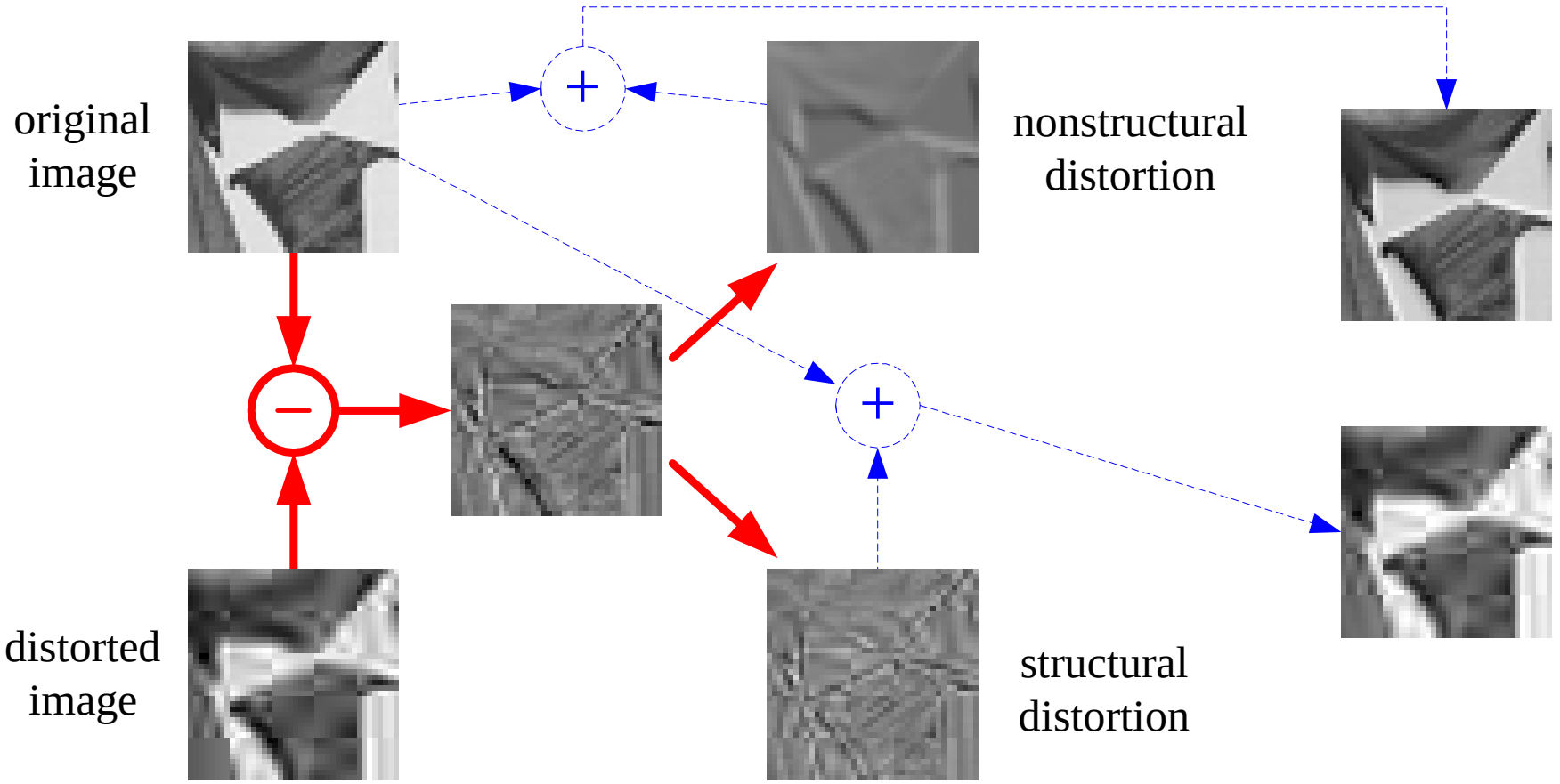
Separation of Structural/Non-Structural Distortions



Separation of Structural/Non-Structural Distortions

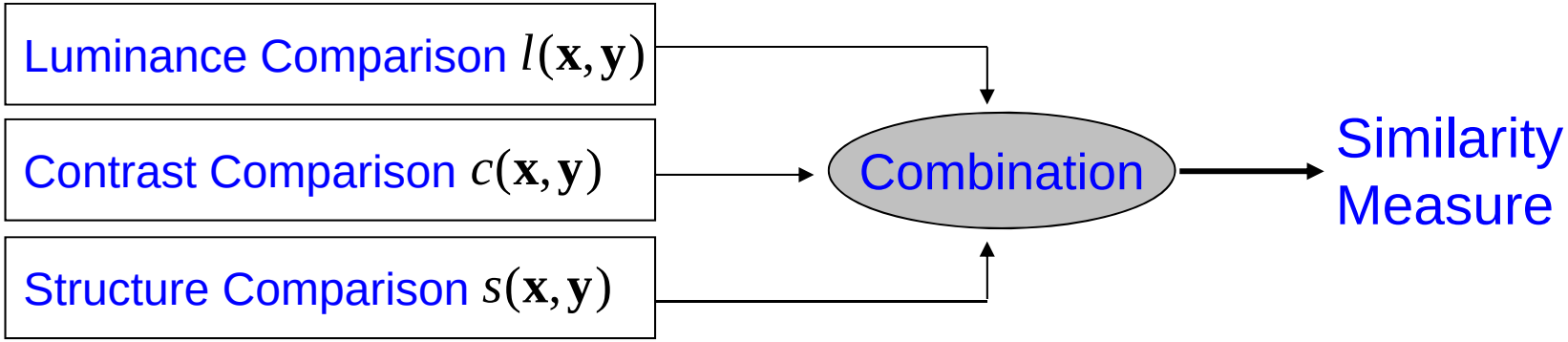


Separation of Structural/Non-Structural Distortions



Structure Similarity Index (SSIM)

For two corresponding local patches $\mathbf{x}$ and $\mathbf{y}$ in two images



Assume that $\mathbf{x}$ and $\mathbf{y}$ are vectorized as

$\mathbf{x} = [x_1, x_2, ..., x_N]$ and $\mathbf{y} = [y_1, y_2, ..., y_N]$

$\mu_x (\mu_y)$ is the mean intensity of $\mathbf{x}$ ($\mathbf{y}$), $\mu_x = \frac{1}{N}\sum_{i=1}^{N} x_i$

$\sigma_x (\sigma_y)$ is the standard deviation of $\mathbf{x}$ ($\mathbf{y}$), $\sigma_x = \left(\frac{1}{N}\sum_{i=1}^{N} (x_i - \mu_x)^2 \right)^{1/2}$

σ_{xy} is the covariance of $\mathbf{x}$ and $\mathbf{y}$, $\sigma_{xy} = \frac{1}{N}\sum_{i=1}^{N} (x_i - \mu_x)(y_i - \mu_y)$

Structure Similarity Index (SSIM)

$$l(\mathbf{x},\mathbf{y}) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}, \quad c(\mathbf{x},\mathbf{y}) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, \quad s(\mathbf{x},\mathbf{y}) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}$$

C_1, C_2, C_3 are fixed constants, and usually set $C_3 = C_2 / 2$

Then, the structure similarity between $\mathbf{x}$ and $\mathbf{y}$ are defined as

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

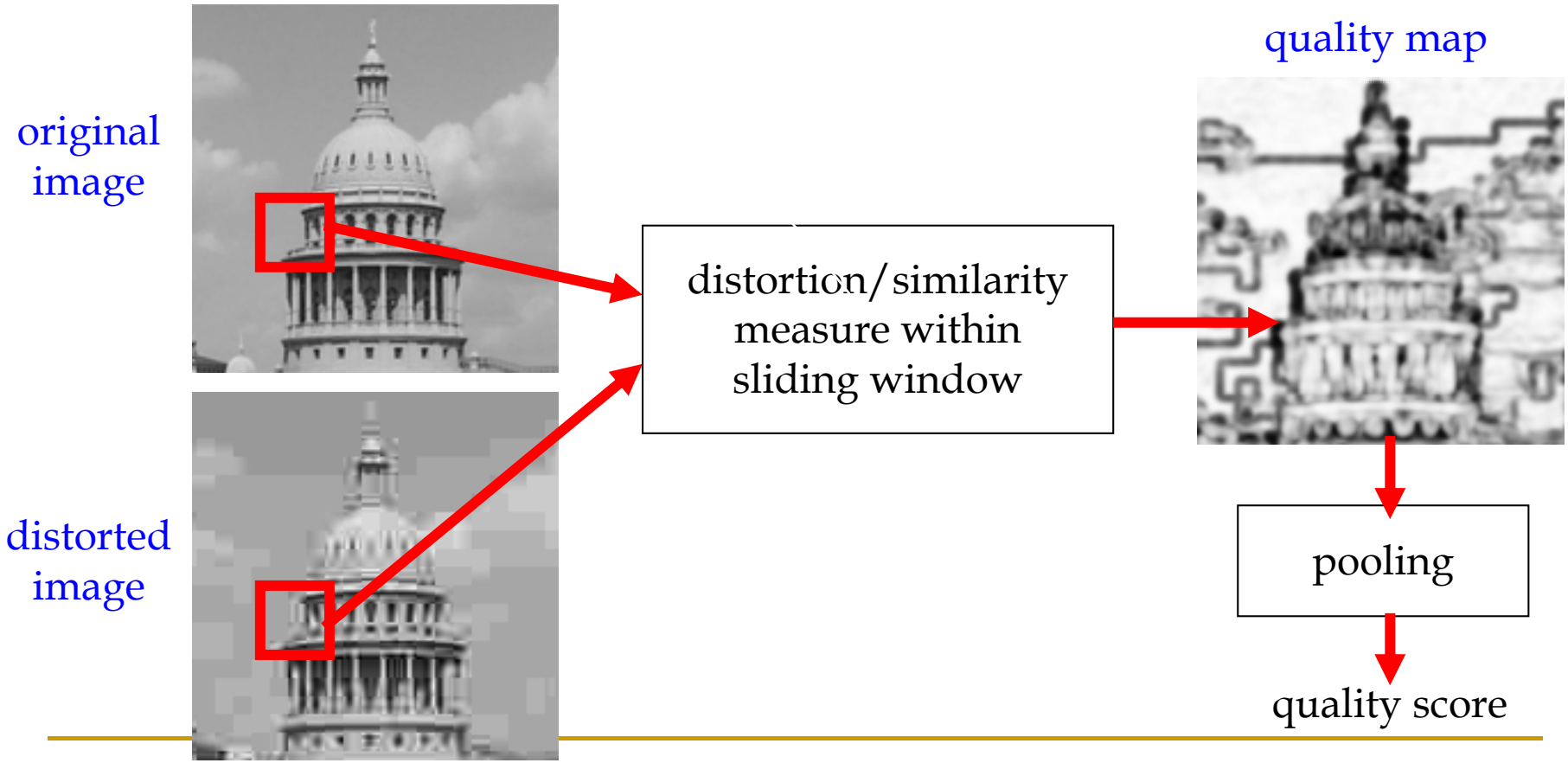
If the image contains M local patches (defined by a sliding window), the overall image quality is

$$\mathrm{SSIM} = \frac{1}{M}\sum_{i=1}^{M} SSIM(\mathbf{x}_i, \mathbf{y}_i)$$

Structure Similarity Index (SSIM)

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

[Wang & Bovik, *IEEE Signal Proc. Letters*, '02]
[Wang *et al.*, *IEEE Trans. Image Proc.*, '04]



Structure Similarity Index (SSIM): Example

$$\begin{bmatrix} 110 & 113 & 113 & 115 \\ 100 & 102 & 102 & 115 \\ 103 & 103 & 108 & 110 \\ 105 & 120 & 106 & 114 \end{bmatrix}$$

reference image

$$\begin{bmatrix} 109 & 112 & 112 & 114 \\ 103 & 104 & 102 & 110 \\ 115 & 103 & 101 & 112 \\ 105 & 125 & 106 & 116 \end{bmatrix}$$

distorted image

Assume that $C_1 = 6.5, C_2 = 58.5$ and the sliding window is 3 by 3

Structure Similarity Index (SSIM): Example

$$\mathbf{x}_1 = \begin{bmatrix} 110 & 113 & 113 & 115 \\ 100 & 102 & 102 & 115 \\ 103 & 103 & 108 & 110 \\ 105 & 120 & 106 & 114 \end{bmatrix} \xrightarrow{\text{distortion}} \mathbf{y}_1 = \begin{bmatrix} 109 & 112 & 112 & 114 \\ 103 & 104 & 102 & 110 \\ 115 & 103 & 101 & 112 \\ 105 & 125 & 106 & 116 \end{bmatrix}$$

reference image distorted image

$$\mathbf{x}_1 = [110\ 113\ 113\ 100\ 102\ 102\ 103\ 103\ 108] \quad \mathbf{y}_1 = [109\ 112\ 112\ 103\ 104\ 102\ 115\ 103\ 101]$$

$$\mu_x = (110 + 113 + \ldots + 108)/9 = 106 \qquad \mu_y = (109 + 112 + \ldots + 101)/9 = 106.8$$

$$\sigma_x = \left((110-106)^2 + (113-106)^2 + \ldots + (108-106)^2\right)/9\Big)^{1/2} = 4.76$$

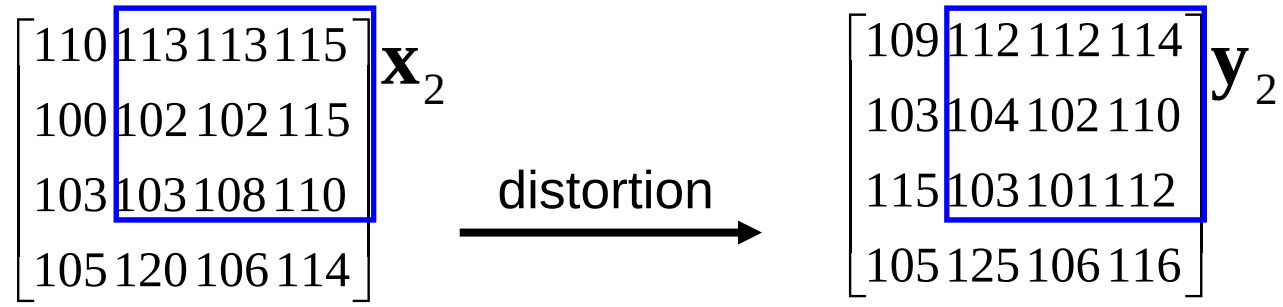
$$\sigma_y = \left((109-106.78)^2 + (112-106.78)^2 + \ldots + (101-106.78)^2\right)/9\Big)^{1/2} = 4.94$$

$$\sigma_{xy} = \left((110-106)*(109-106.78) + (113-106)*(112-106.78) + \ldots + (108-106)*(101-106.78)\right)/9$$

$$= 12.22$$

$$SSIM(\mathbf{x}_1, \mathbf{y}_1) = \frac{(2*106*106.78 + 6.5)*(2*12.22 + 58.5)}{(106^2 + 106.78^2 + 6.5)*(4.76^2 + 4.94^2 + 58.5)} = 0.7857$$

Structure Similarity Index (SSIM): Example



$$\mathbf{x}_2 = [113\ 113\ 115\ 102\ 102\ 115\ 103\ 108\ 110] \quad \mathbf{y}_2 = [112\ 112\ 114\ 104\ 102\ 110\ 103\ 101\ 112]$$

$$\mu_x = (113 + 113 + \ldots + 110)/9 = 107.89 \qquad \mu_y = (112 + 112 + \ldots + 112)/9 = 107.78$$

$$\sigma_x = \left((113-107.89)^2 + (113-107.89)^2 + \ldots + (110-107.89)^2\right)/9\Big)^{1/2} = 4.82$$

$$\sigma_y = \left((112-107.78)^2 + (112-107.78)^2 + \ldots + (112-107.78)^2\right)/9\Big)^{1/2} = 4.87$$

$$\sigma_{xy} = \big((113-107.89)*(112-107.78) + (113-107.89)*(112-107.78) + \ldots + (110-107.89)*(112-107.78)\big)/9$$

$$= 18.75$$

$$SSIM(\mathbf{x}_2, \mathbf{y}_2) = \frac{(2*107.89*107.78 + 6.5)*(2*18.72 + 58.5)}{(107.89^2 + 107.78^2 + 6.5)*(4.82^2 + 4.87^2 + 58.5)} = 0.9105$$

Structure Similarity Index (SSIM): Example

$$\mathbf{x}_3 = \begin{bmatrix} 110 & 113 & 113 & 115 \\ 100 & 102 & 102 & 115 \\ 103 & 103 & 108 & 110 \\ 105 & 120 & 106 & 114 \end{bmatrix} \xrightarrow{\text{distortion}} \mathbf{y}_3 = \begin{bmatrix} 109 & 112 & 112 & 114 \\ 103 & 104 & 102 & 110 \\ 115 & 103 & 101 & 112 \\ 105 & 125 & 106 & 116 \end{bmatrix}$$

reference image distorted image

$$\mathbf{x}_3 = [100\ 102\ 102\ 103\ 103\ 108\ 105\ 120\ 106] \quad \mathbf{y}_3 = [103\ 104\ 102\ 115\ 103\ 101\ 105\ 125\ 106]$$

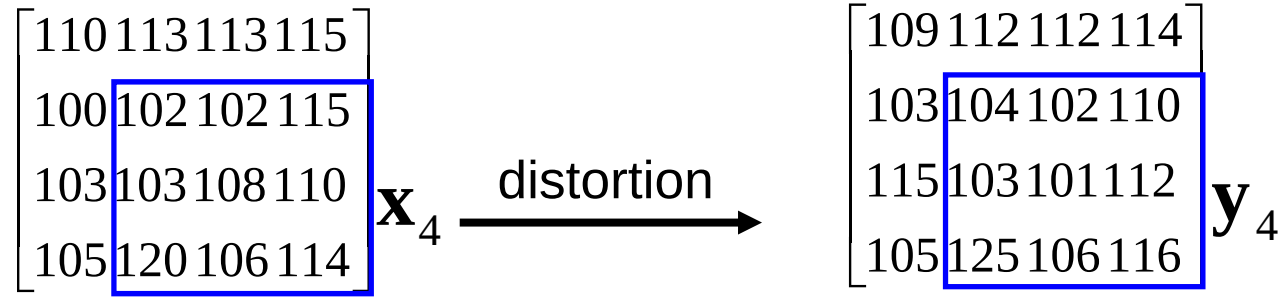
$$\mu_x = 105.44 \qquad \mu_y = 107.11$$

$$\sigma_x = 5.62 \qquad \sigma_y = 7.42$$

$$\sigma_{xy} = 31.84$$

$$SSIM(\mathbf{x}_3, \mathbf{y}_3) = 0.8421$$

Structure Similarity Index (SSIM): Example



reference image

distorted image

$$\mathbf{x}_4 = [102\ 102\ 115\ 103\ 108\ 110\ 120\ 106\ 114] \quad \mathbf{y}_4 = [104\ 102\ 110\ 103\ 101\ 112\ 125\ 106\ 116]$$

$$\mu_x = 107.78 \qquad \mu_y = 108.78$$

$$\sigma_x = 5.71 \qquad \sigma_y = 7.44$$

$$\sigma_{xy} = 38.28$$

$$SSIM(\mathbf{x}_4, \mathbf{y}_4) = 0.9225$$

Structure Similarity Index (SSIM): Example

$$\begin{bmatrix} 110 & 113 & 113 & 115 \\ 100 & 102 & 102 & 115 \\ 103 & 103 & 108 & 110 \\ 105 & 120 & 106 & 114 \end{bmatrix} \xrightarrow{\text{distortion}} \begin{bmatrix} 109 & 112 & 112 & 114 \\ 103 & 104 & 102 & 110 \\ 115 & 103 & 101 & 112 \\ 105 & 125 & 106 & 116 \end{bmatrix}$$

reference image distorted image

The overall similarity of the reference image and the distorted image is

$$\text{SSIM} = \frac{1}{4}\sum_{i=1}^{4} SSIM(\mathbf{x}_i, \mathbf{y}_i)$$

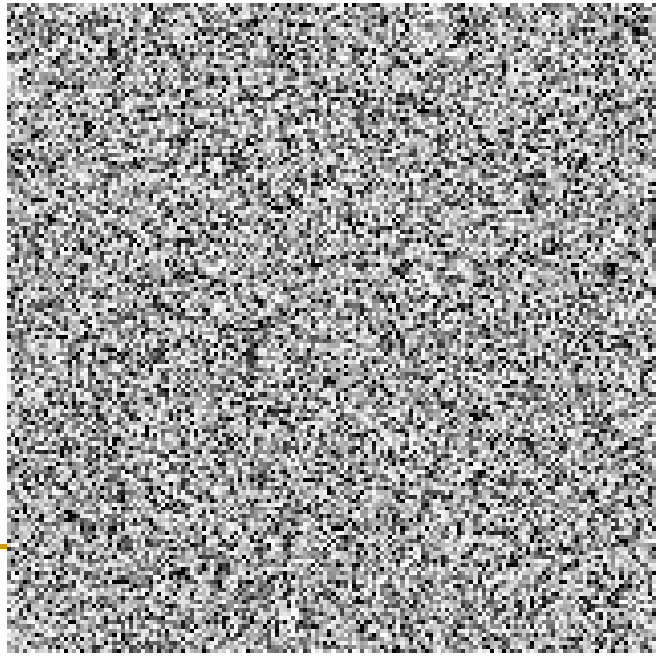
$$= \frac{1}{4}\left(0.7857 + 0.9105 + 0.8421 + 0.9225\right) = 0.8652$$

Gaussian noise corrupted image

original image

SSIM index map

absolute error map

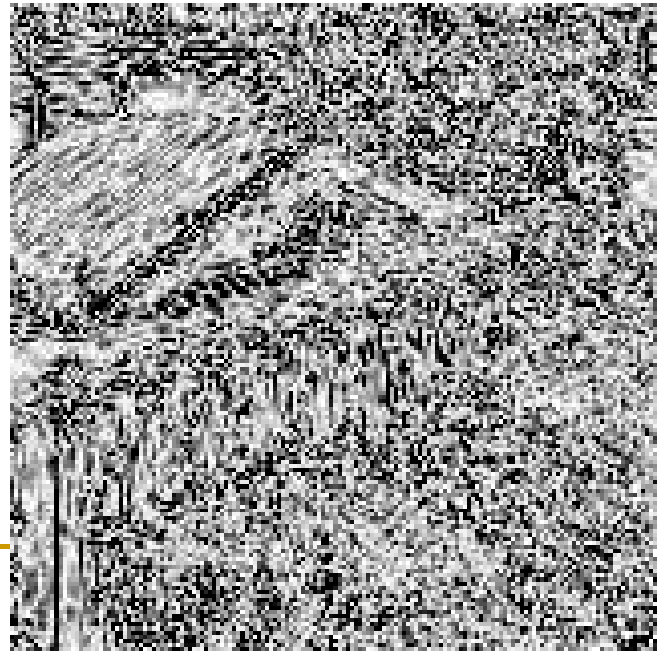


JPEG2000 compressed image

original image

SSIM index map

absolute error map

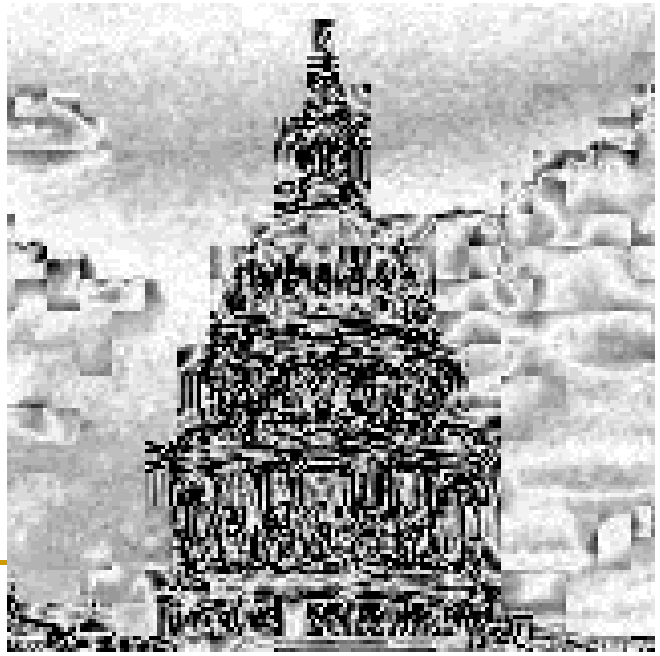


JPEG compressed image

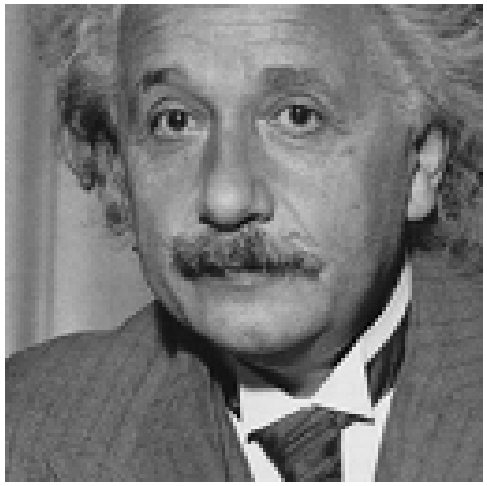
original image

SSIM index map

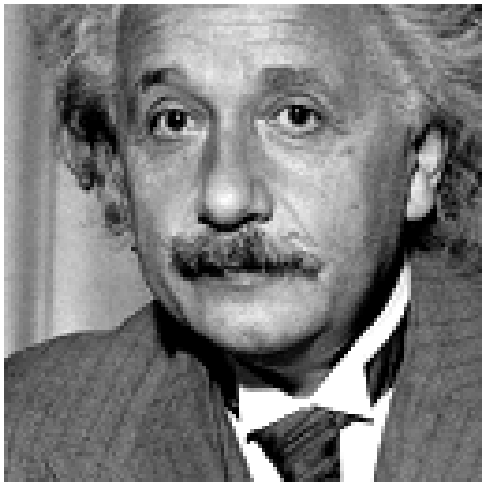
absolute error map



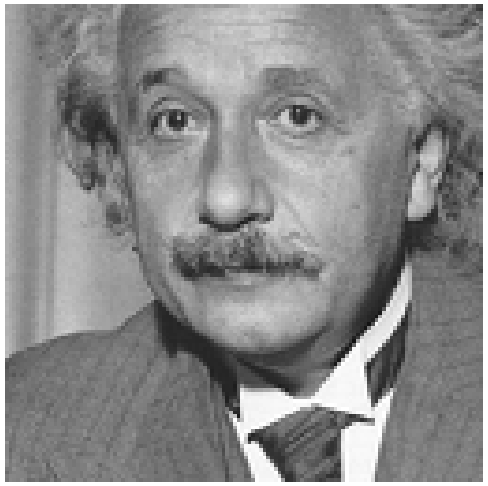
original Image



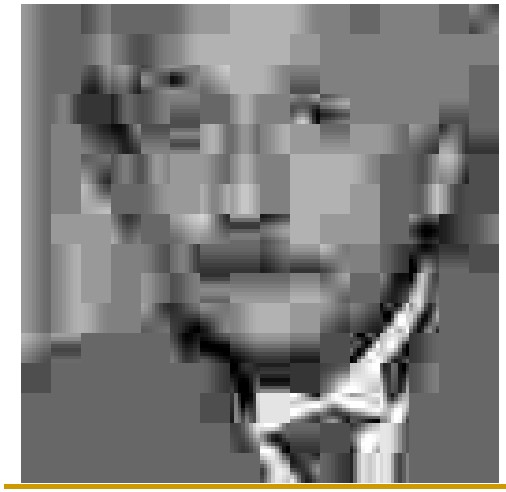
MSE=0, SSIM=1



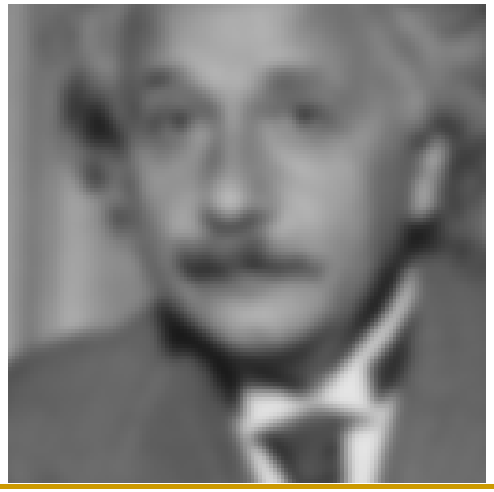
MSE=309, SSIM=0.928



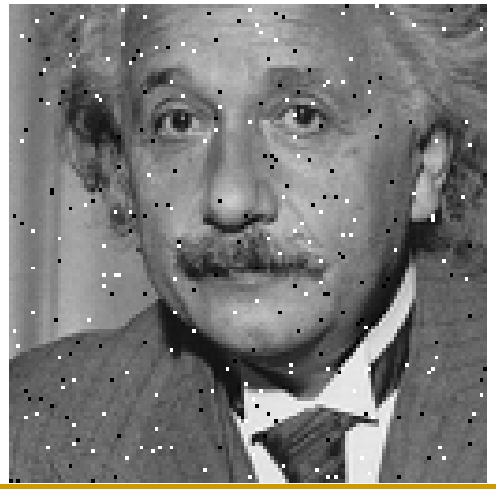
MSE=309, SSIM=0.987



MSE=309, SSIM=0.580

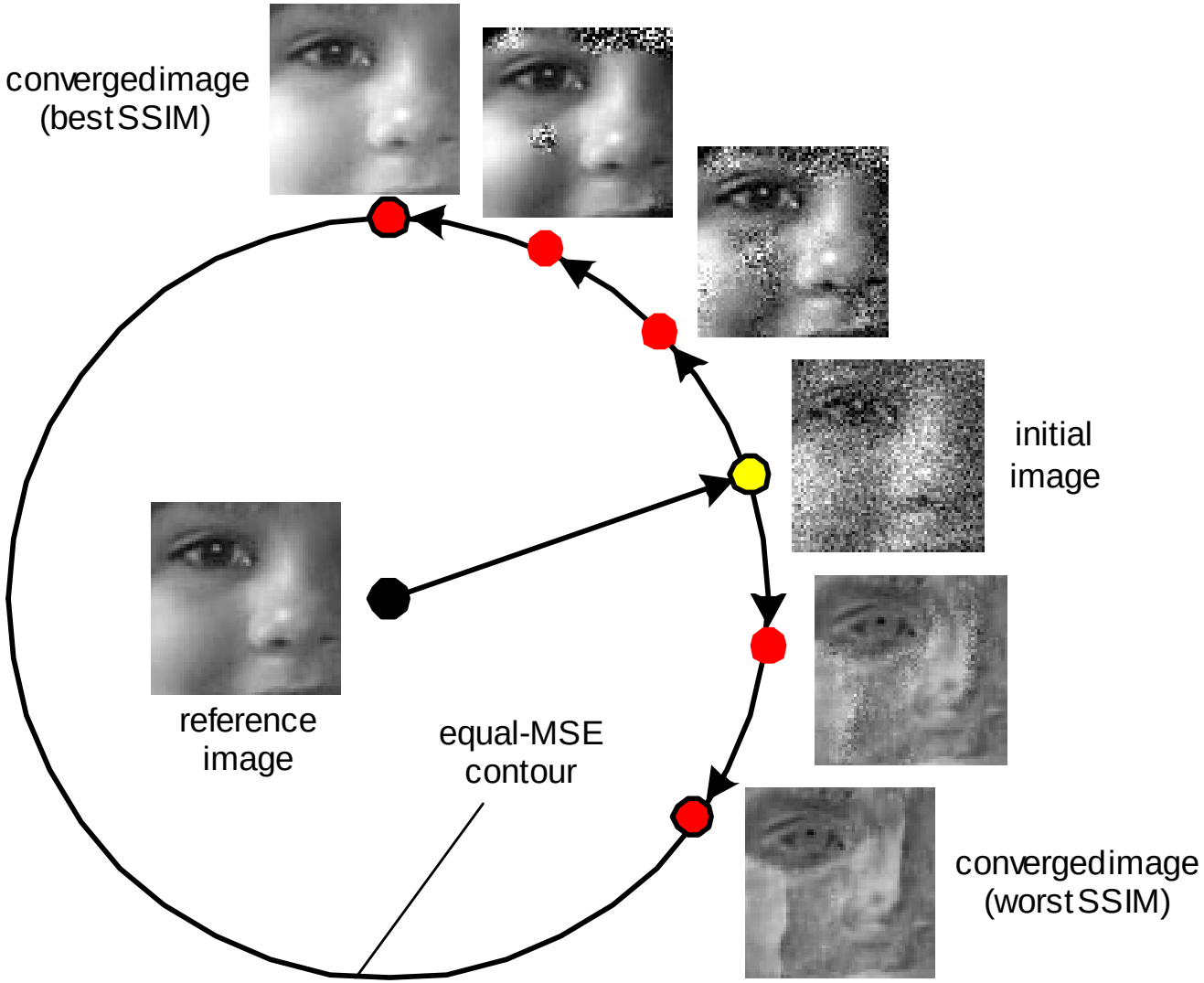


MSE=309, SSIM=0.641

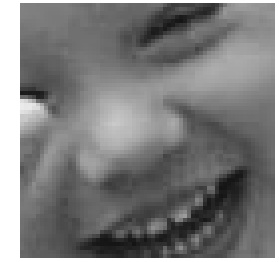


MSE=309, SSIM=0.730

MSE vs. SSIM

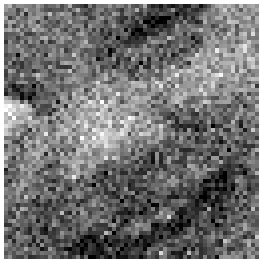

converged image
(best SSIM)
initial
image
reference
image
equal-MSE
contour
converged image
(worst SSIM)

MSE vs. SSIM



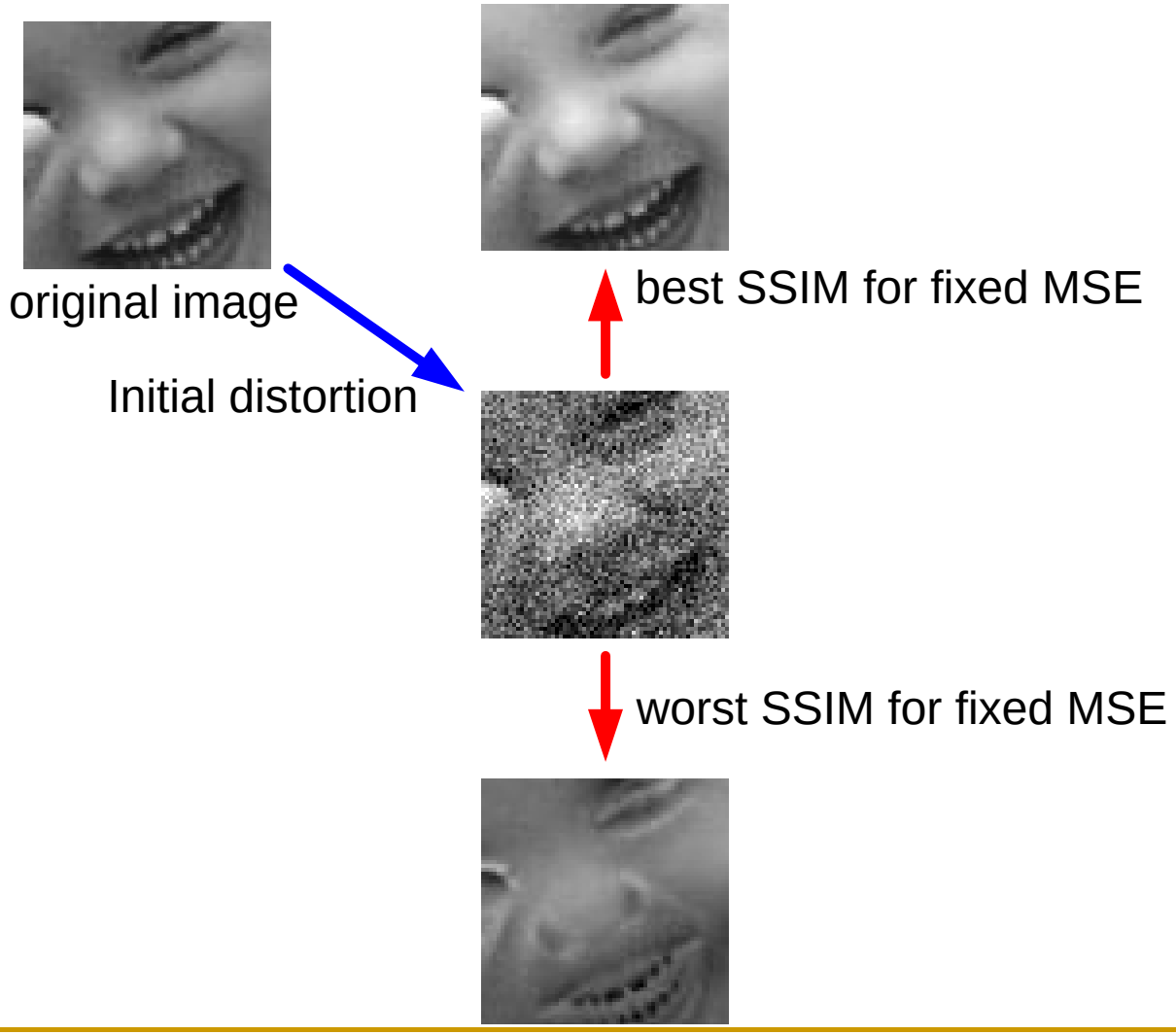
original image

Initial distortion

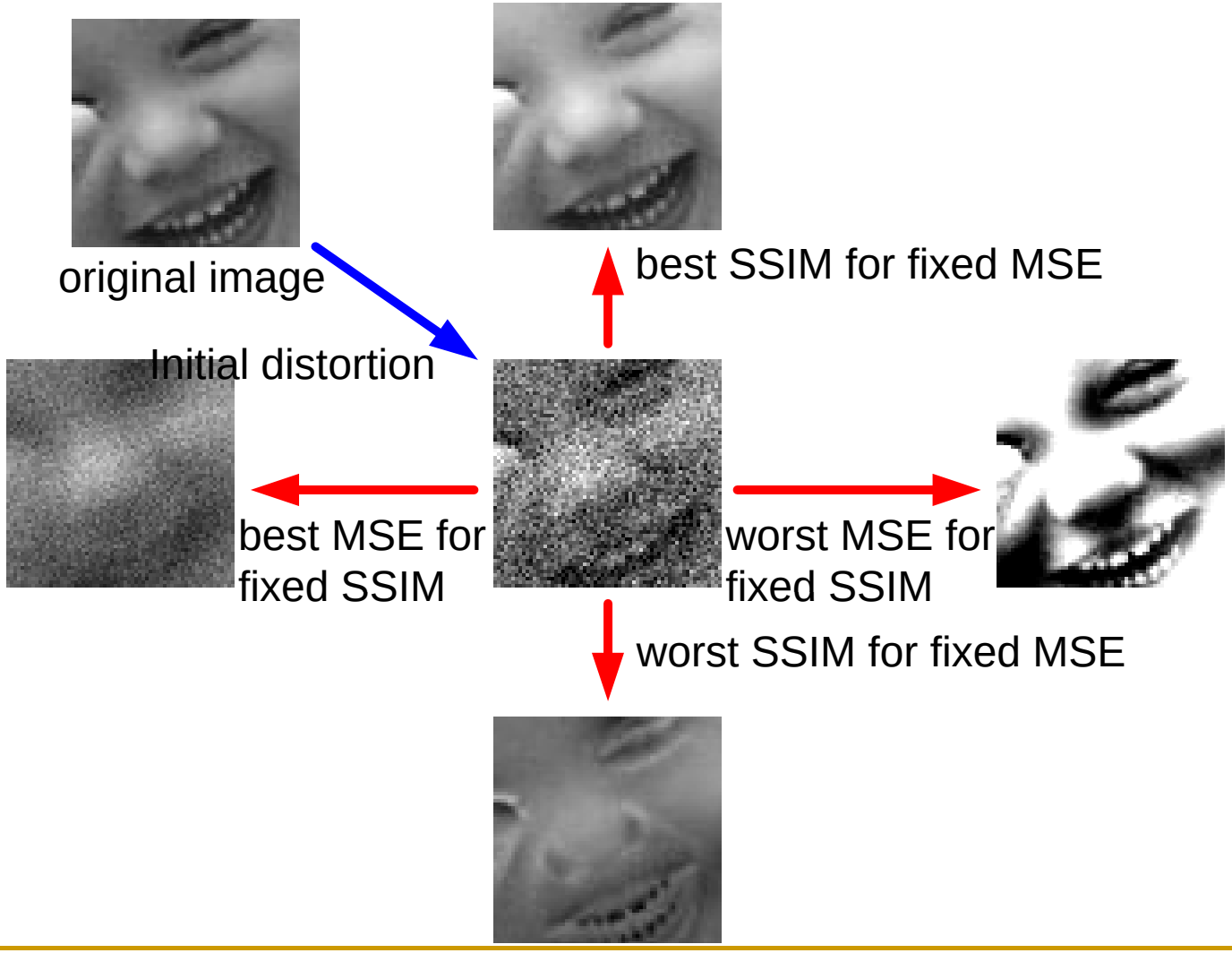


Initial image

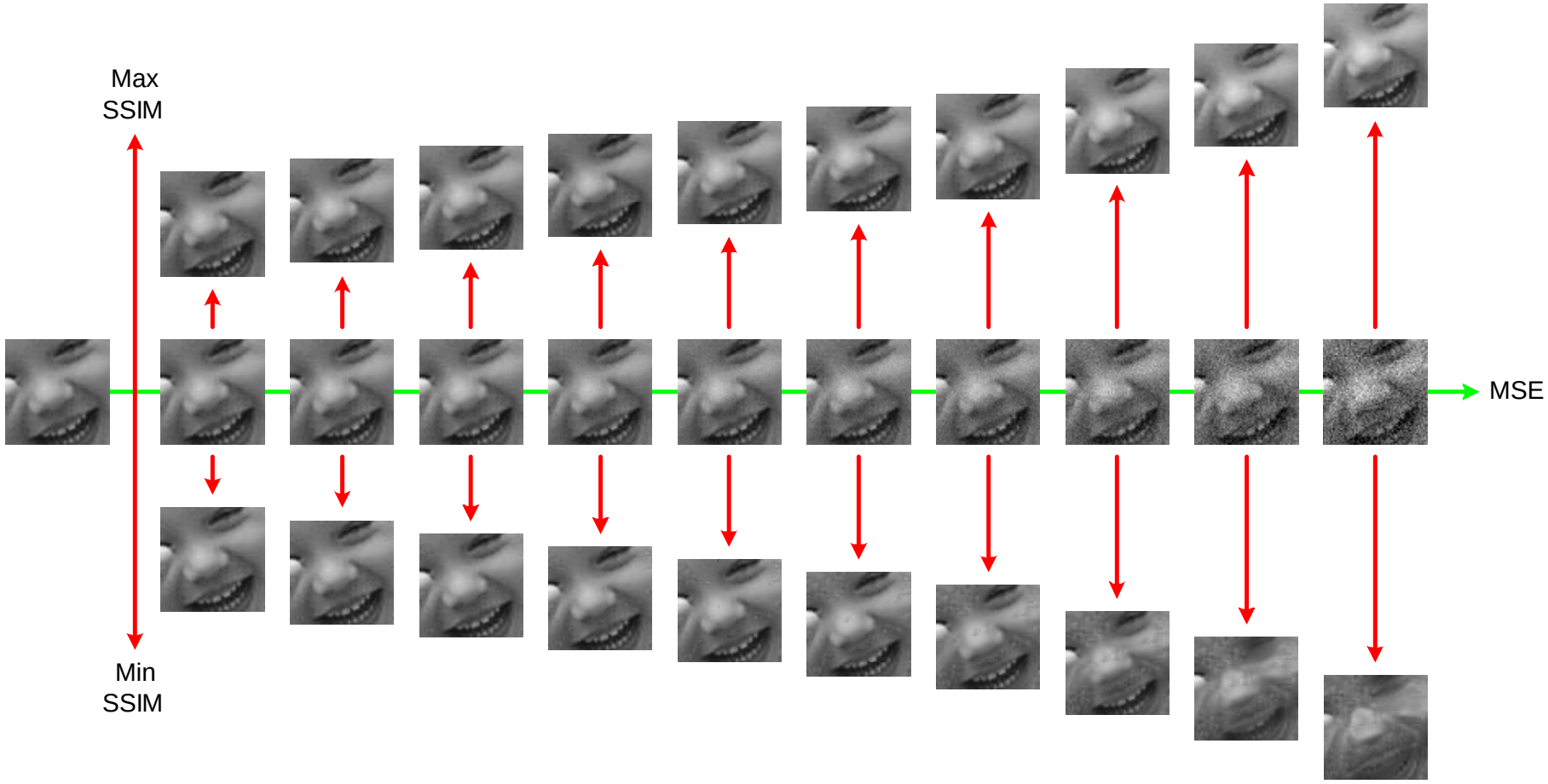
MSE vs. SSIM


original image
Initial distortion
best SSIM for fixed MSE
worst SSIM for fixed MSE

MSE vs. SSIM


original image
Initial distortion
best SSIM for fixed MSE
best MSE for fixed SSIM
worst MSE for fixed SSIM
worst SSIM for fixed MSE

MSE vs. SSIM



Using SSIM

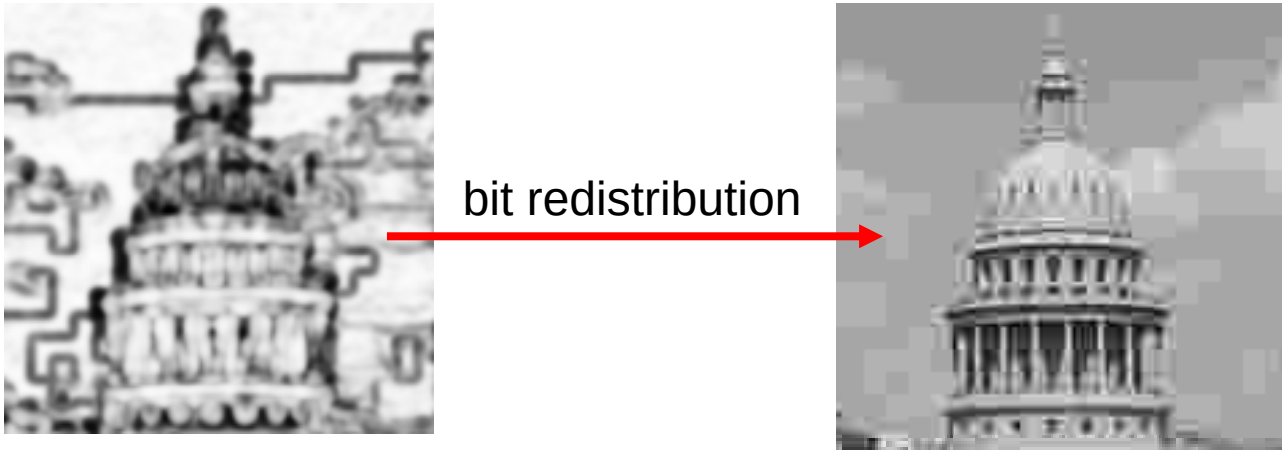
- **Image/video coding and communications**
- **Watermarking/data hiding**
- **Image denoising**
- **Image enhancement**
- **Image/video hashing**
- **Image fusion**
- **Superresolution/interpolation**
- **Image/texture synthesis**
- **Image halftoning**
- **Vision processor design**
- **Display design**
- **Contrast equalization for LCD**
- ……

Outline

- Why “Perceptual”?
  - Motivation
  - Overview
- Perceptual Image Quality Assessment
  - Mean squared error’s deficiency
  - Error visibility methods
  - Structural similarity methods
- Perceptual Image Processing
  - Image compression
  - ……

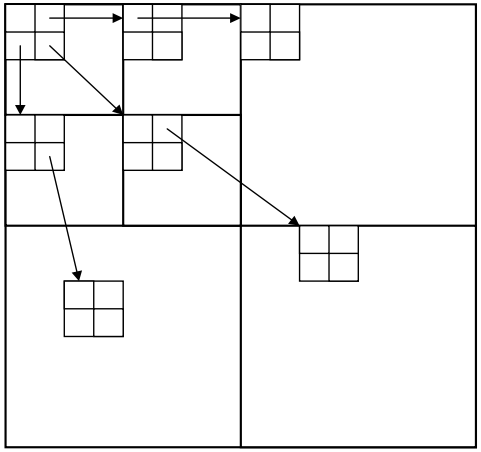
Perceptual Image Compression

- General Idea
  - Transform image signal into “perceptually uniform” space
- Implementations
  - Perceptual weighting + uniform quantization
  - Equivalently, perceptually adaptive quantization
  - Net effect: bit redistribution, perceptually

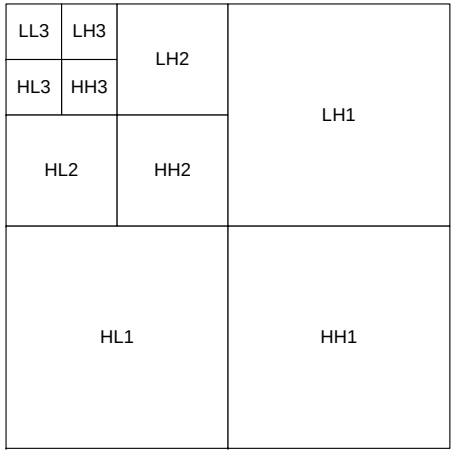


Perceptual Image Compression

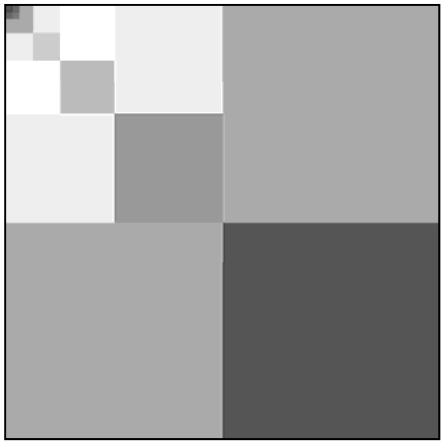
- Frequency Weighting
  - JPEG quantization table; JPEG2000 subband weighting



wavelet tree



wavelet subbands



subband weighting

- Masking
  - JPEG2000 neighborhood/self masking

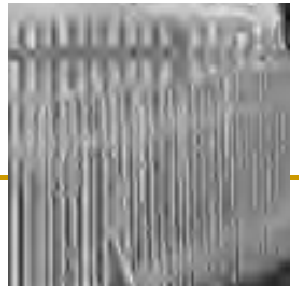
Perceptual Image Compression

original image

SPIHT, 0.2bits/pixel

SSIM map (SPIHT)

New, 0.2bits/pixel



SSIM map (new)

[Wang, Li & Shang '07]

References

- Zhou Wang and Alan C. Bovik. *Modern Image Quality Assessment*. Synthesis Lectures on Image, Video & Multimedia Processing, Morgan & Claypool Publishers, 2006.
- L. Zhang, L. Zhang, X. Mou, D. Zhang, “FSIM: A feature similarity index for image quality assessment,” Image Processing, IEEE Transactions on 20 (8), 2378-2386, 2011.