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
Define Perceptual IQA Measures

Optimize IP Systems & Algorithms “Perceptually”

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
  — NR & RR: wide open research topic

  \[ PSNR = 10 \log_{10} \frac{L^2}{MSE} \]

• IQA is Difficult
• 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

Key features
- **Channel decomposition** → linear frequency/orientation transforms
- **Frequency weighting** → contrast sensitivity function
- **Masking** → intra/inter channel interaction

\[
E = \left[ \sum_{l} \sum_{k} |e_{l,k}|^\beta \right]^{1/\beta}
\]
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]
New Paradigm: Structural Similarity

**Philosophy**

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

HVS is highly adapted for this purpose

Estimate **structural information change**

<table>
<thead>
<tr>
<th>Classical philosophy</th>
<th>New philosophy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bottom-up</td>
<td>Top-down</td>
</tr>
<tr>
<td>Predict Error Visibility</td>
<td>Predict Structural Distortion</td>
</tr>
</tbody>
</table>

- How to define structural information?
- How to separate structural/nonstructural information?
Separation of Structural/nonstructural Distortion

original image

+ 

original image

- 

distorted image

+ 

distorted image

+ 

distorted image
Separation of Structural/nonstructural Distortion
Separation of Structural/nonstructural Distortion

original image -> + -> distorted image

- -> structural distortion

+ -> nonstructural distortion
Separation of Structural/nonstructural Distortion

original image

nonstructural distortion

distorted image

structural distortion
Adaptive Linear System
Adaptive Linear System

Original image

Distorted image

\[ = \Delta c_1 \cdot + \Delta c_2 \cdot + \cdots + \Delta c_K \cdot \]

\[ = \Delta c_{K+1} \cdot + \Delta c_{K+2} \cdot + \cdots + \Delta c_M \cdot \]
Adaptive Linear System

overcomplete, adaptive basis in the space of all images

[Wang & Simoncelli, ICIP '05, submitted]
Structural Similarity (SSIM) Index in Image Space

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

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

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

\[ SSIM(x, y) = l(x, y) \cdot c(x, y) \cdot s(x, y) \]

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

Minkowski (MSE)  component-weighted  magnitude-weighted

magnitude and component-weighted  SSIM
Gaussian noise corrupted 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

<table>
<thead>
<tr>
<th>Dataset</th>
<th>JP2(1)</th>
<th>JP2(2)</th>
<th>JPG(1)</th>
<th>JPG(2)</th>
<th>Noise</th>
<th>Blur</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td># of images</td>
<td>87</td>
<td>82</td>
<td>87</td>
<td>88</td>
<td>145</td>
<td>145</td>
<td>145</td>
</tr>
<tr>
<td>PSNR</td>
<td>0.934</td>
<td>0.895</td>
<td>0.902</td>
<td>0.914</td>
<td><strong>0.987</strong></td>
<td>0.774</td>
<td>0.881</td>
</tr>
<tr>
<td>SSIM</td>
<td><strong>0.968</strong></td>
<td>0.967</td>
<td><strong>0.965</strong></td>
<td><strong>0.986</strong></td>
<td>0.971</td>
<td>0.936</td>
<td>0.944</td>
</tr>
</tbody>
</table>

**Graphs:**
- PSNR vs. MOS for different datasets and image types, with fitting curves.
- MSSIM vs. MOS, showing performance metrics across different conditions.
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]
Extensions of SSIM (1)

- Color image quality assessment
  [Toet & Lucassen., Displays, ’03]

- Video quality assessment

- 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

\[
\text{SSIM}(x, y) = \frac{2|\sum c_x \cdot c_y^*| + C}{\sum |c_x|^2 + \sum |c_y|^2 + C}
\]

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

[Wang & Simoncelli, ICASSP '05]
Image Matching without Registration

**Standard patterns:** 10 images

\[
\begin{array}{cccccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 0 \\
\end{array}
\]

**Database:** 2430 images

\[
\begin{array}{cccccccccccc}
3 & 1 & 4 & 5 & 0 & 9 & 3 & 8 & 9 & 4 & 7 & 8 & 4 & 1 & 3 & 1 \\
5 & 4 & 6 & 9 & 3 & 2 & 5 & 9 & 6 & 8 & 2 & 0 & 6 & 4 & 7 & 1 \\
8 & 3 & 8 & 7 & 6 & 5 & 8 & 9 & 8 & 7 & 8 & 7 & 0 & 2 & 2 & 0 \\
5 & 2 & 9 & 4 & 6 & 8 & 2 & 4 & 6 & 1 & 1 & 0 & 2 & 1 & 8 & 6 \\
0 & 5 & 6 & 6 & 0 & 7 & 5 & 7 & 6 & 2 & 6 & 8 & 4 & 2 & 5 & 1 \\
8 & 6 & 6 & 1 & 3 & 5 & 7 & 2 & 0 & 5 & 0 & 3 & 2 & 7 & 8 & 9 \\
\end{array}
\]

**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!