# 2018 Computer Vision

## Image Quality Assessment

GuoZemin
2018.06.11

# Overview

1. **Introduction to Image Quality Assessment**

- Image characteristic
- Image quality index
- What is Image Quality Assessment
- Application of Image Quality Assessment

2. Datasets and The metrics of Image Quality Assessment algorithm

- The introduction to the datasets
- The metrics of Image Quality Assessment algorithm

# Overview

3 **Introduction to methods**

- Subjective Assessment
- Objective Assessment

4 Deep learning framework -- PyTorch

- Brief introduction to the PyTorch
- PyTorch vs TensorFlow

# PSNR vs MOS

- PSNR: Peak Signal-to-Noise Ratio
- MOS: Mean Opinion Score

# What is the MSE(Mean Square Error)?

Suppose that x = $\{x_i \mid i = 1,2,\ldots,N\}$ and y = $\{y_i \mid i = 1,2,\ldots,N\}$ are two finite-length, discrete signals (e.g., visual images), where N is the number of signal samples (pixels, if the signals are images) and $x_i$ and $y_i$ are the values of the i th samples in x and y, respectively. The MSE between the signals is

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

In the literature of image processing, MSE is often converted into a peak signal-to-noise ratio (PSNR) measure

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

https://ieeexplore.ieee.org/abstract/document/4775883

# The MSE has many attractive feature:

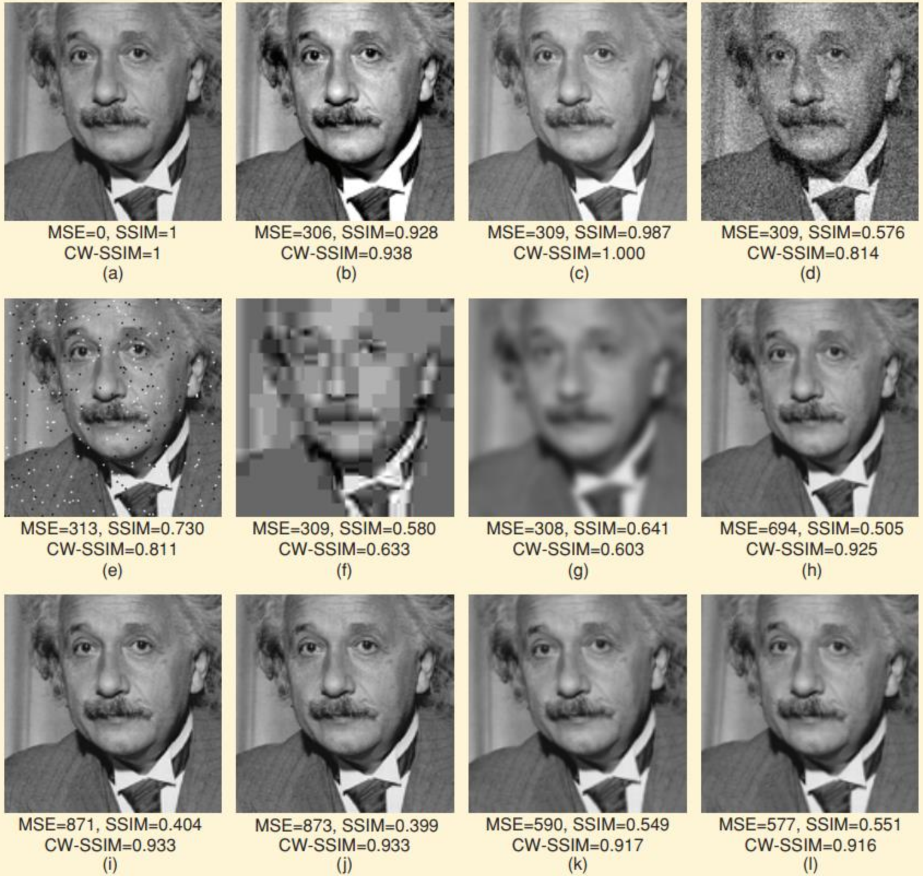
1. It is simple. It is parameter free and inexpensive to compute, with a complexity of  per sample is low. It is also memoryless-the squared error can be evaluated at each sample, independent of other samples.

# The MSE has many attractive feature:

2. It has a clear physical meaning – it is the natural way to define the energy of the error signal.
3. The MSE is an excellent metric in the context of optimization. The MSE possesses the very satisfying properties of convexity, symmetry, and differentiability.

$$PSNR = 10log_{10}\frac{L^2}{MSE} \stackrel{?}{=}$$

https://ieeexplore.ieee.org/abstract/document/4775883



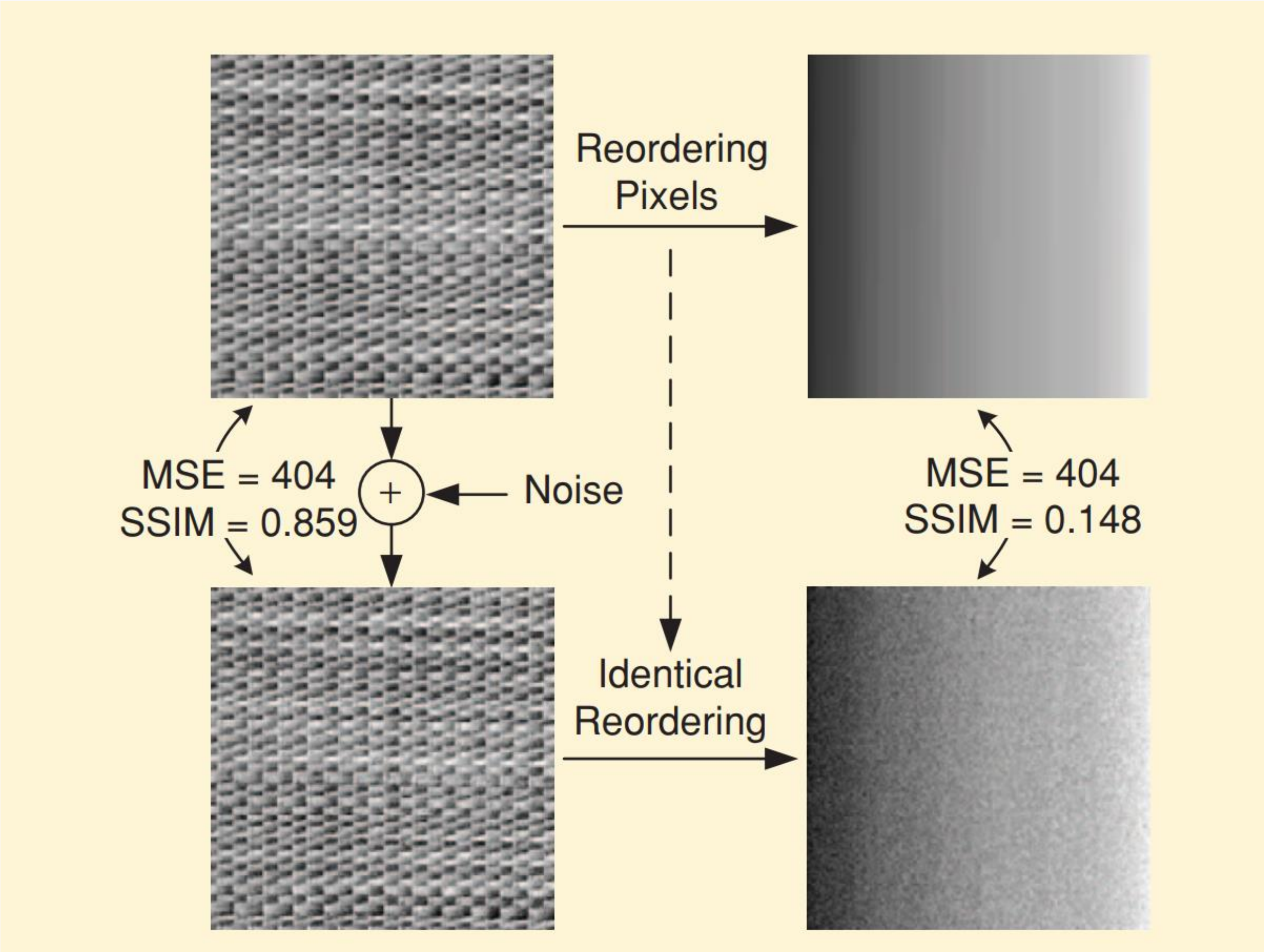

**[FIG2] Comparison of image fidelity measures for "Einstein" image altered with different types of distortions. (a) Reference image. (b) Mean contrast stretch. (c) Luminance shift. (d) Gaussian noise contamination. (e) Impulsive noise contamination. (f) JPEG compression. (g) Blurring. (h) Spatial scaling (zooming out). (i) Spatial shift (to the right). (j) Spatial shift (to the left). (k) Rotation (counter-clockwise). (l) Rotation (clockwise).**

https://ieeexplore.ieee.org/abstract/document/4775883

# Implicit assumptions when using the MSE:

1. Signal fidelity is independent of temporal or spatial relationships between the samples of the original signal. In other words, if the original and distorted signals are randomly re-ordered in the same way, then the MSE between them will be unchanged.

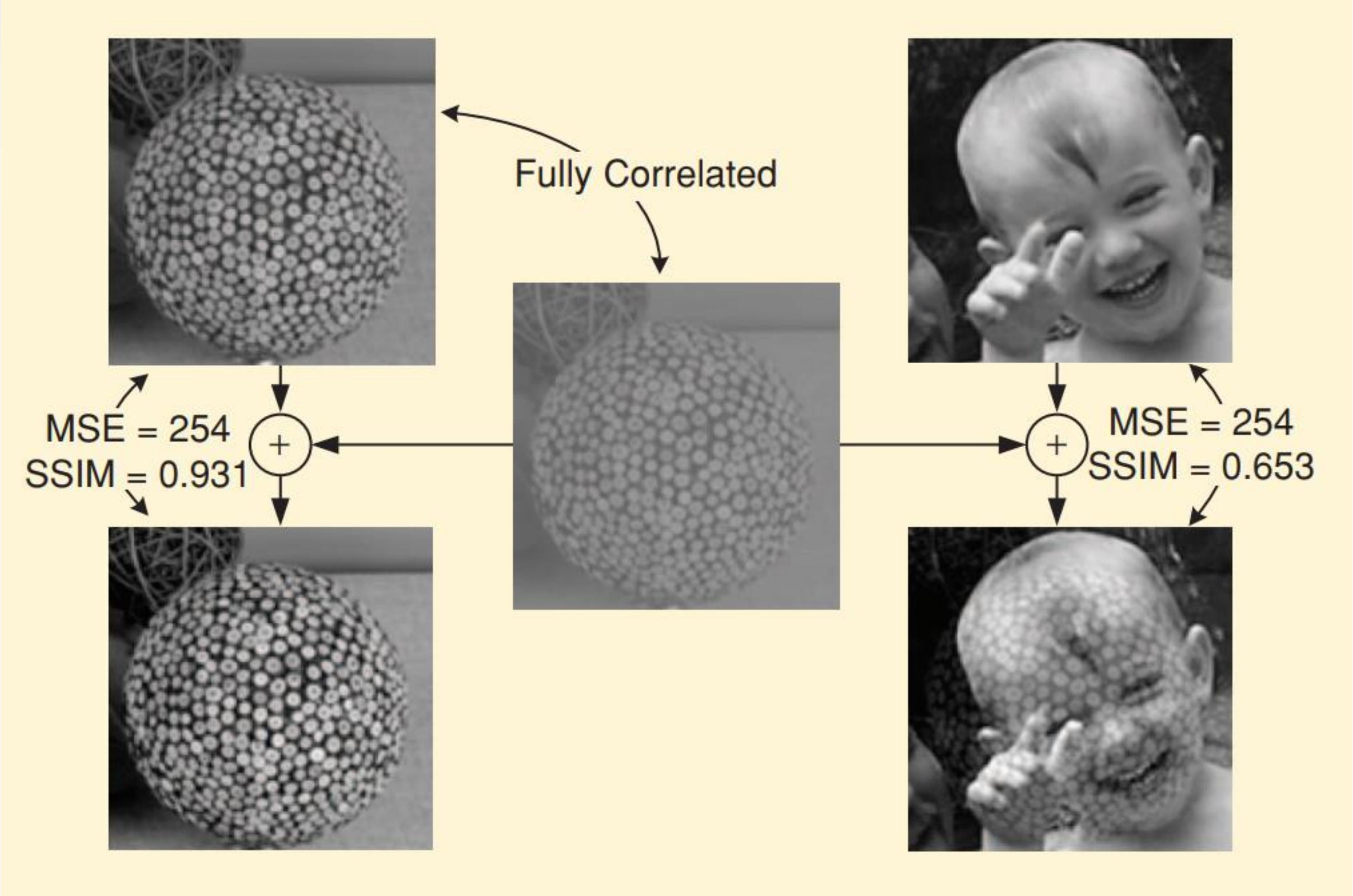
https://ieeexplore.ieee.org/abstract/document/4775883


Reordering
Pixels
MSE = 404
SSIM = 0.859
Noise
MSE = 404
SSIM = 0.148
Identical
Reordering

# Implicit assumptions when using the MSE:

2. Signal fidelity is independent of any relationship between the original signal and the error signal. For a given error signal, the MSE remains unchanged, regardless of which original signal it is added to.

https://ieeexplore.ieee.org/abstract/document/4775883

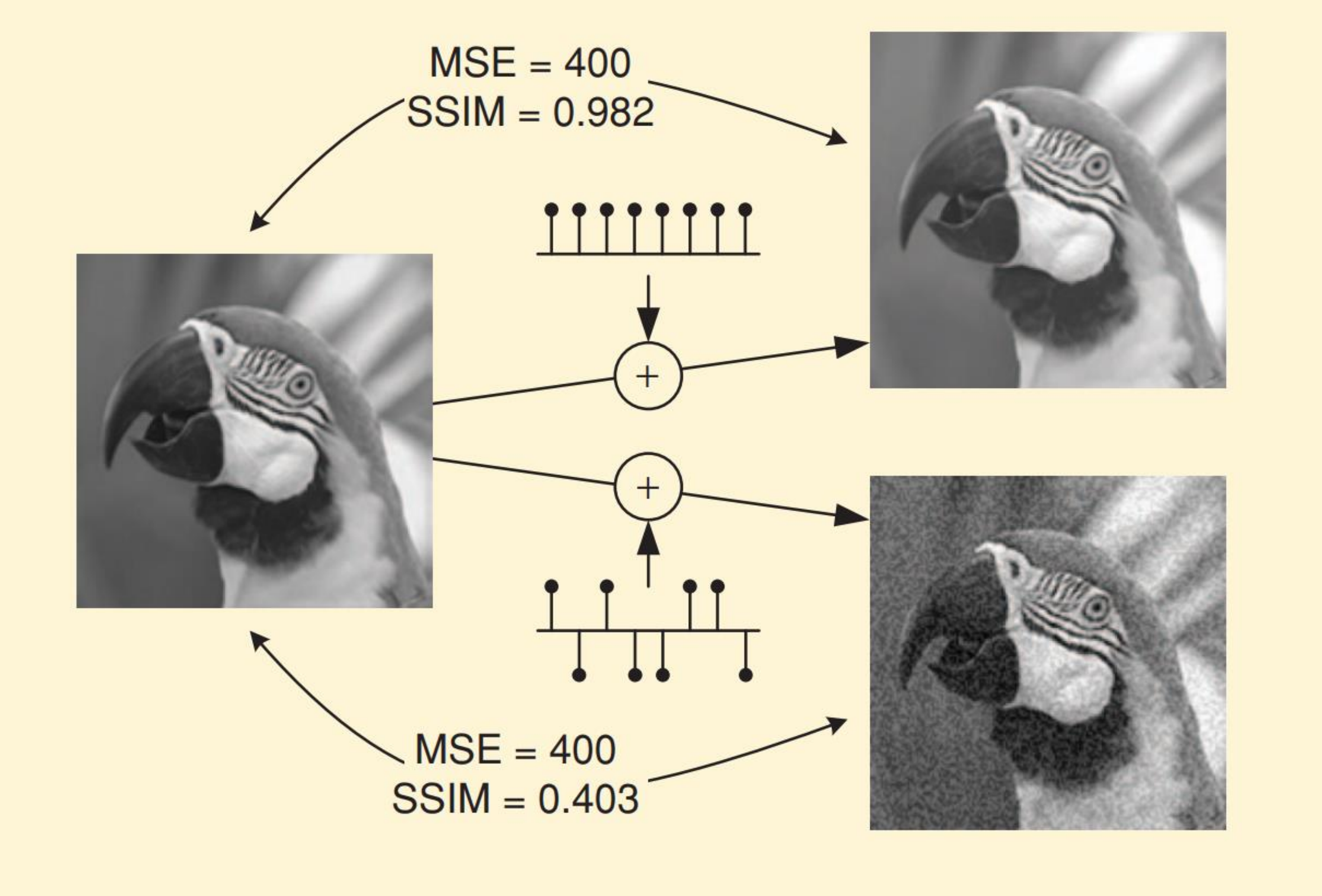



https://ieeexplore.ieee.org/abstract/document/4775883

# Implicit assumptions when using the MSE:

3. Signal fidelity is independent of the signs of the error signal samples.

https://ieeexplore.ieee.org/abstract/document/4775883

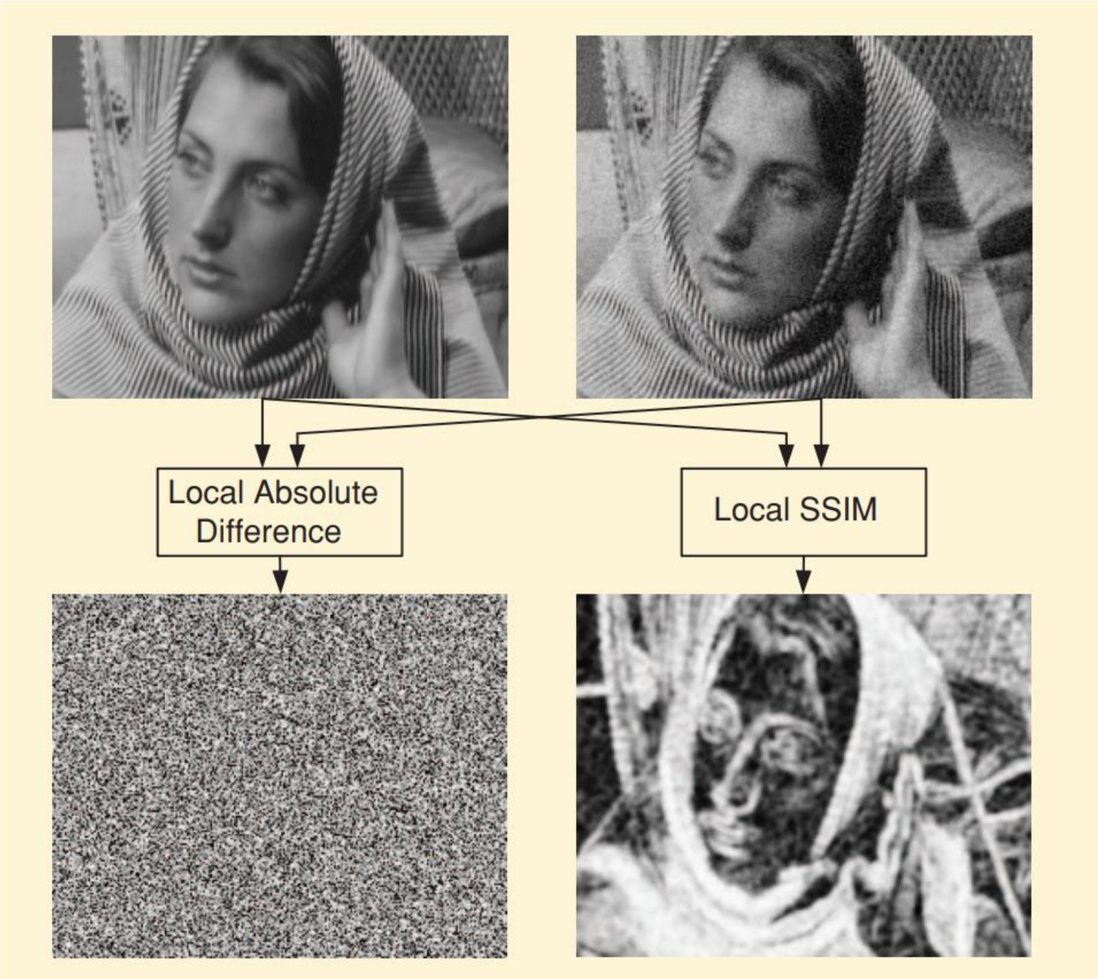



https://ieeexplore.ieee.org/abstract/document/4775883

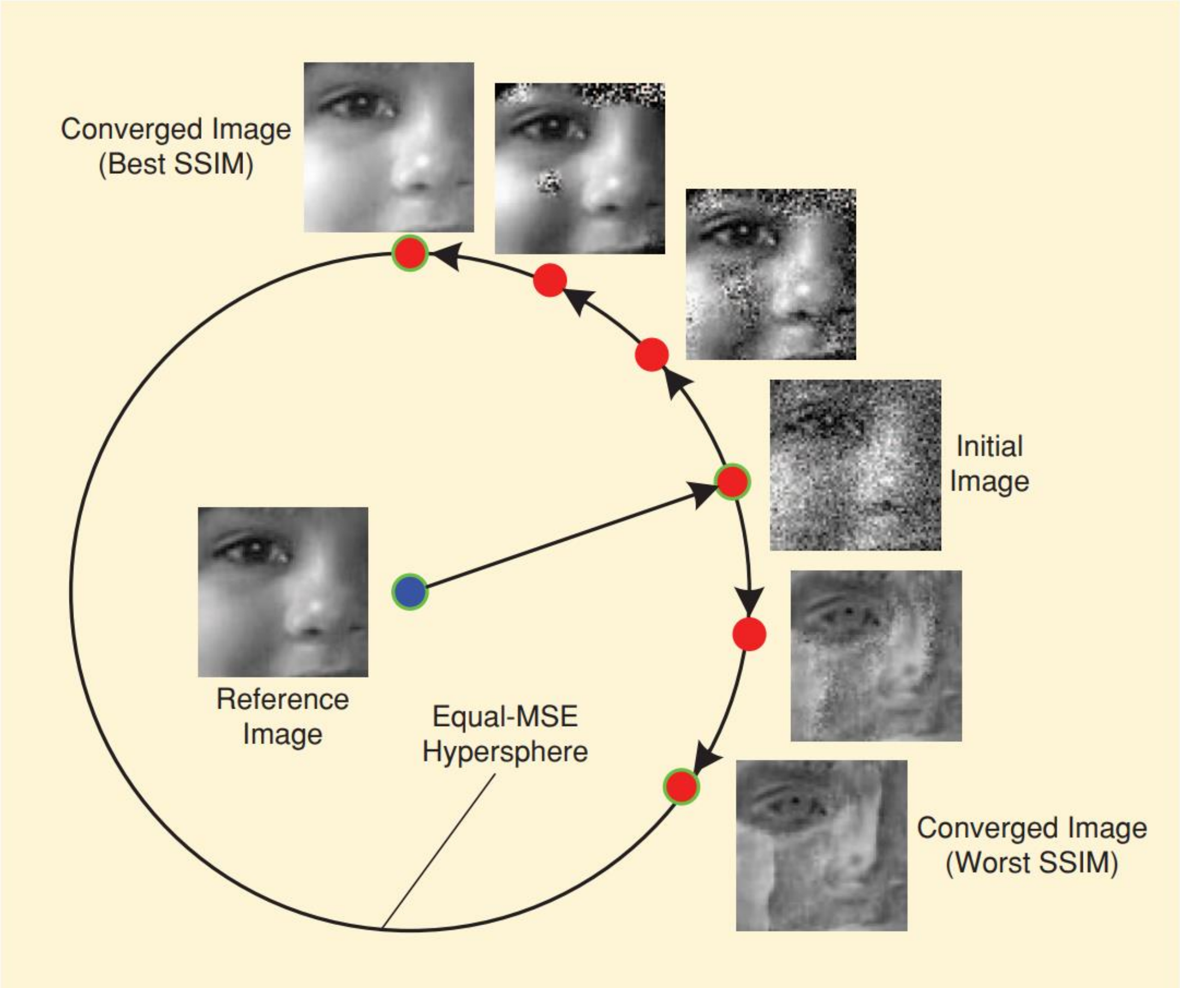
# Implicit assumptions when using the MSE:

4. All signal samples are equally important to signal fidelity.

https://ieeexplore.ieee.org/abstract/document/4775883




https://ieeexplore.ieee.org/abstract/document/4775883



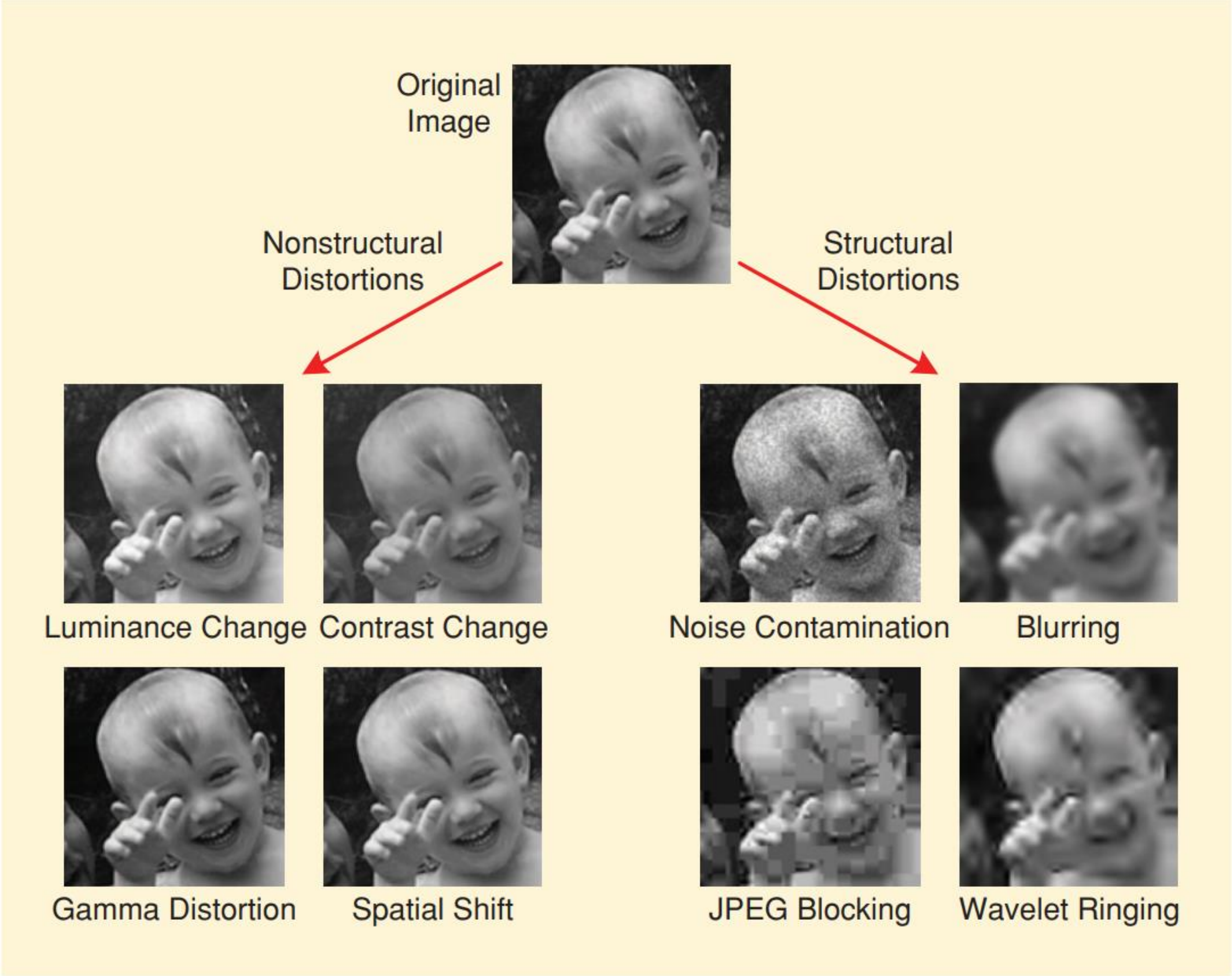

https://ieeexplore.ieee.org/abstract/document/4775883

# Image Quality Assessment: From Error Visibility to Structural Similarity

Zhou Wang, Member, IEEE, Alan Conrad Bovik, Fellow, IEEE, Hamid Rahim Sheikh, Student Member, IEEE, and Eero P. Simoncelli, Senior Member, IEEE



https://ieeexplore.ieee.org/document/1284395



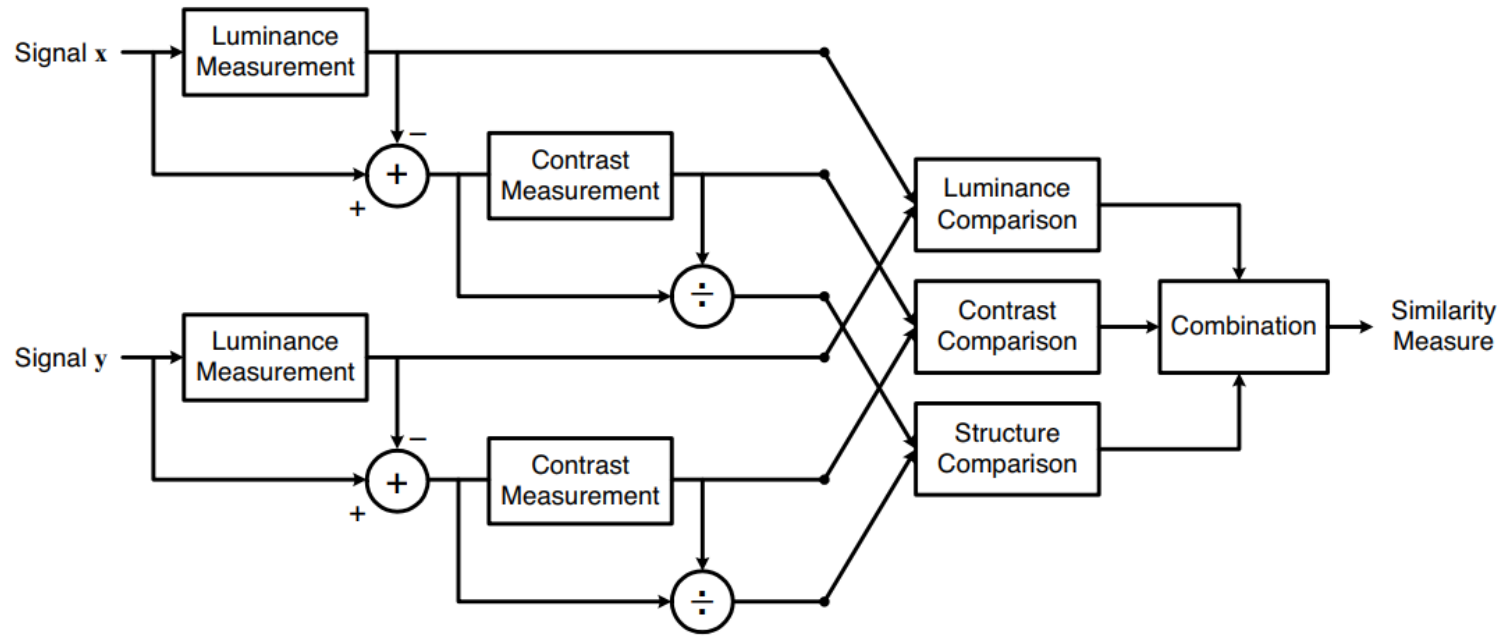

https://ieeexplore.ieee.org/abstract/document/4775883

# Structural Similarity Index(SSIM)

- Luminance
- Contrast
- structure



https://ieeexplore.ieee.org/abstract/document/1284395/




https://ieeexplore.ieee.org/abstract/document/1284395/

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

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

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

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



https://ieeexplore.ieee.org/abstract/document/1284395/

1

# Introduction to Image Quality Assessment

## 1.1 Image characteristic

- Colour feature
- Texture feature
- Shape feature
- Spatial relationship feature

点击此处添加标题

在此输入相关文字，在此输入相关文字，在此输入相关文字在此输入相关文字，在此输入相关文字，在此输入相关文字， 在此输入相关文字，在此输入相关文字，在此输入相关文字在此输入相关文字，在此输入相关文字，在此输入相关文字。



https://blog.csdn.net/qq_23291783/article/details/52036729

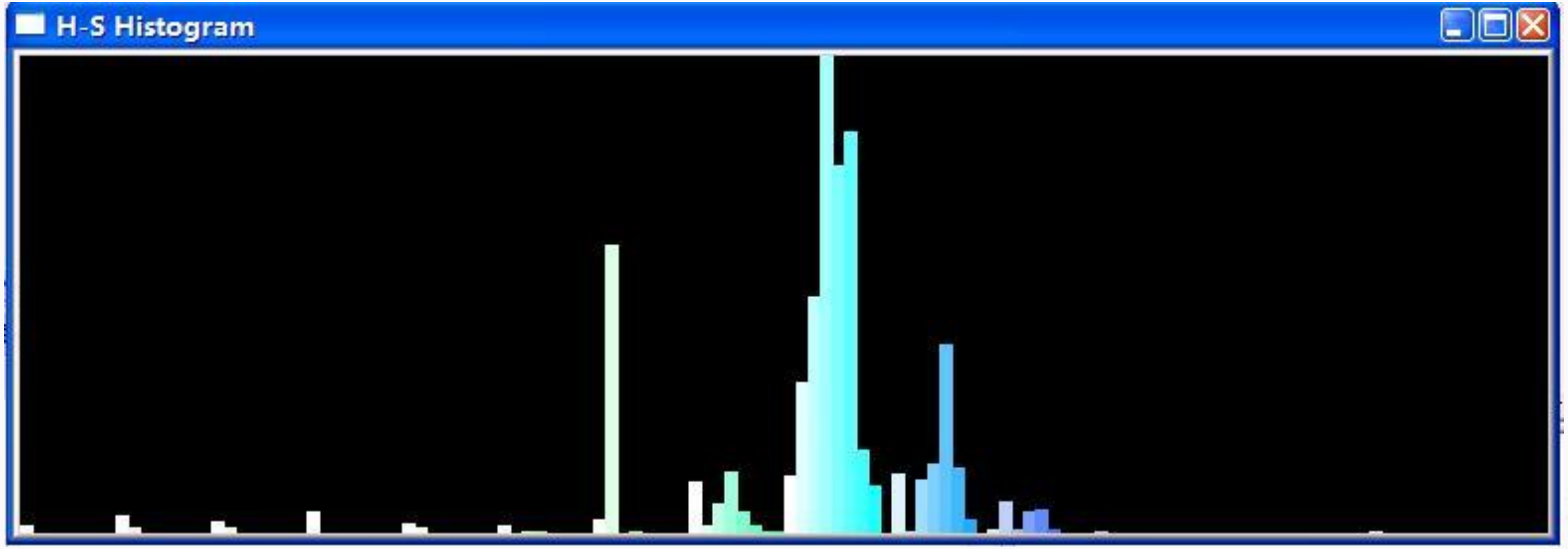
- Colour feature
  - Feature: colour feature is a global feature
  - Methods of expressing color feature
    - **Colour histogram**
    - Colour set
    - Colour moment
    - Colour coherence vector
    - Colour correlogram

## 1.1 Image characteristic

Examples of colour histogram:

# 1.1 Image characteristic

Examples of colour histogram:




https://wenku.baidu.com/view/56481ebef121dd36a32d826b.html

## 1.1 Image characteristic

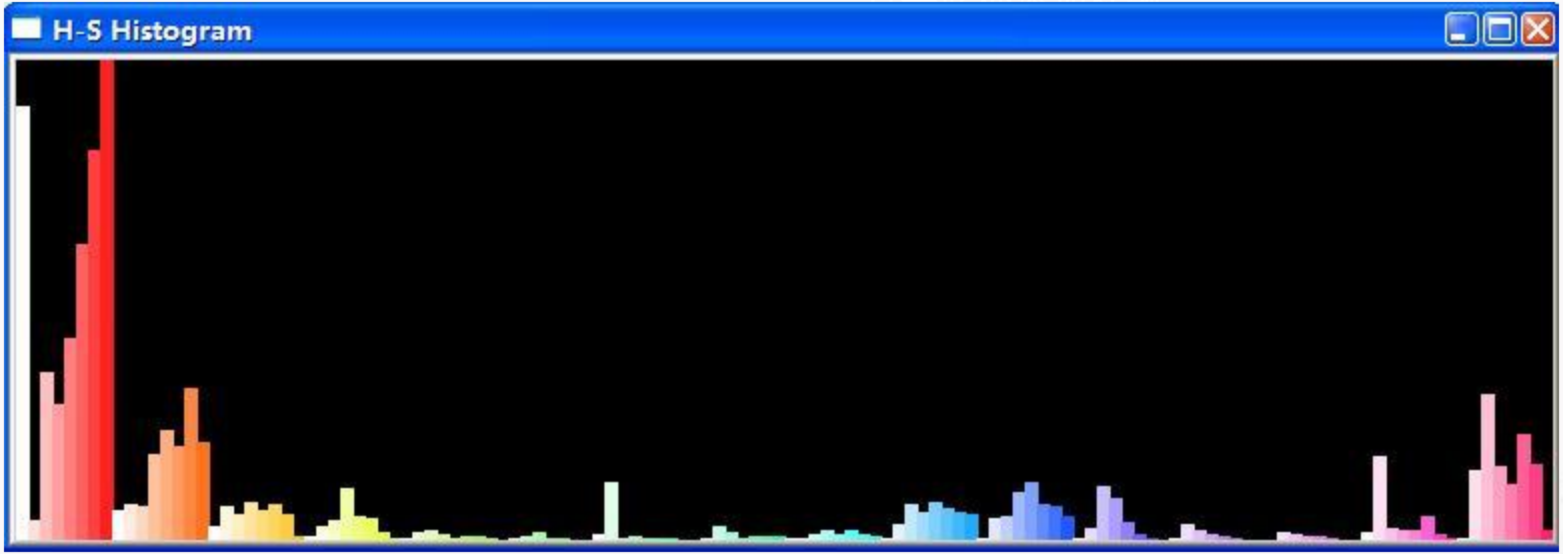
Examples of colour histogram:



https://wenku.baidu.com/view/56481ebef121dd36a32d826b.html

# 1.1 Image characteristic

Examples of colour histogram:



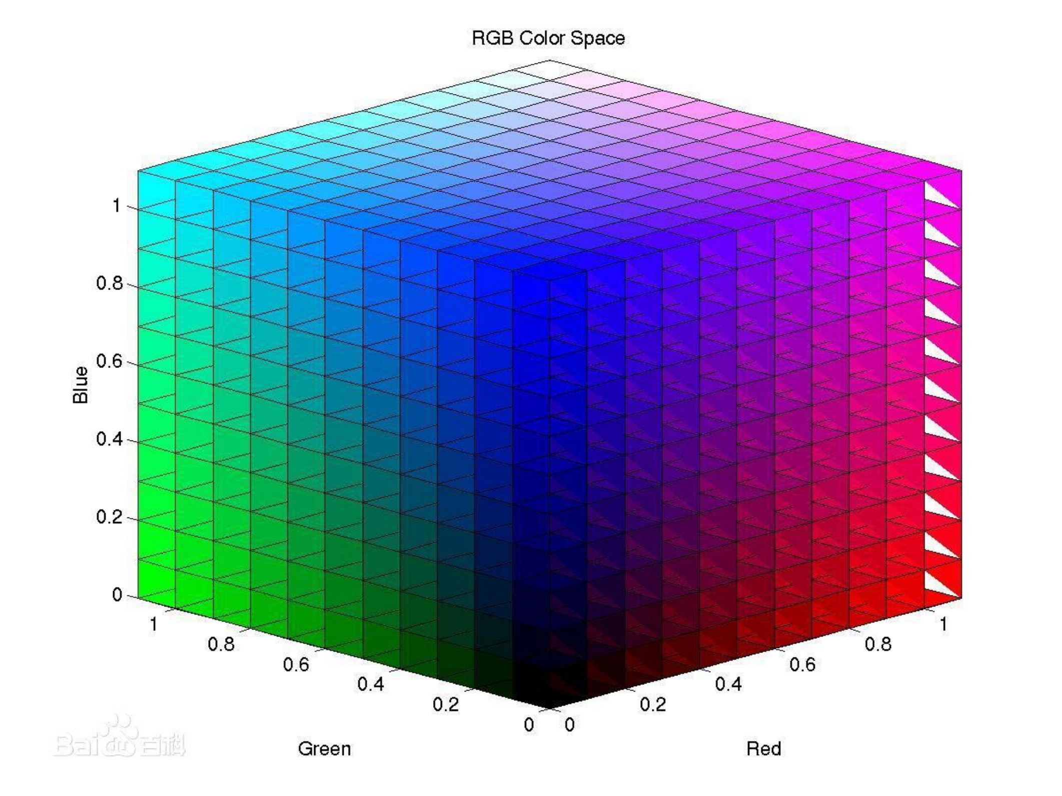



https://wenku.baidu.com/view/56481ebef121dd36a32d826b.html

# 1.1 Image characteristic

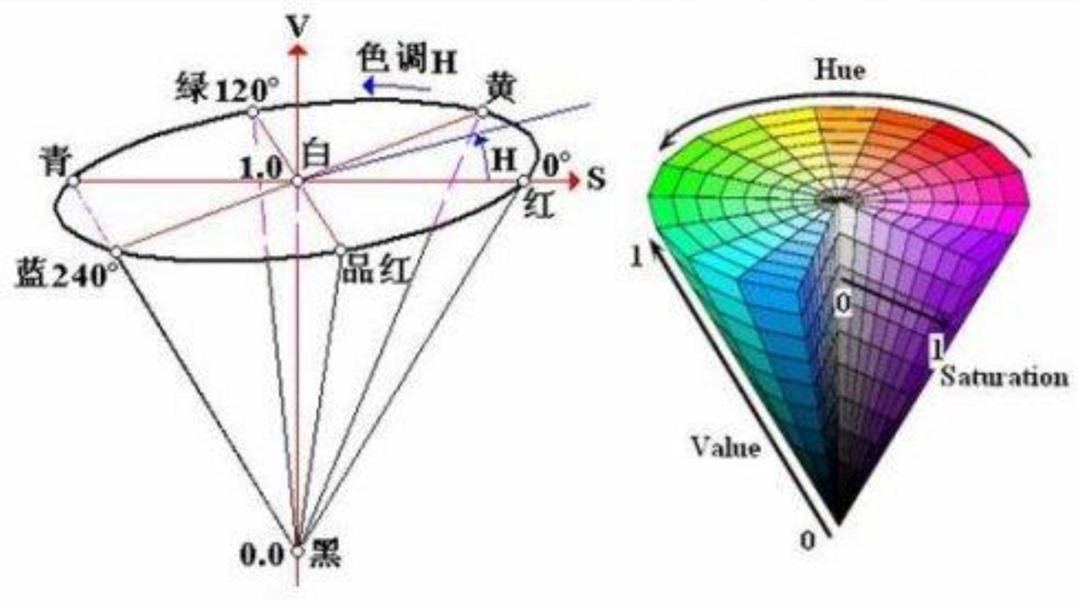
- Colour space: RGB

$$F = r[R] + r[G] + r[B]$$




https://www.cnblogs.com/justkong/p/6570914.html

# Colour space: HSV(hue, saturation, value)

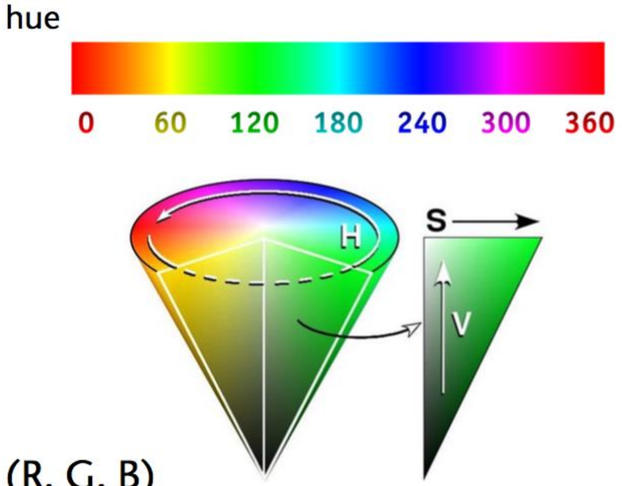





https://www.cnblogs.com/justkong/p/6570914.html

# 1.1 Image characteristic

- colour cone
  - $H$ = hue / colour in degrees $\in [0,360]$
  - $S$ = saturation $\in [0,1]$
  - $V$ = value $\in [0,1]$




- conversion RGB $\rightarrow$ HSV
  - $V = \max = \max(R, G, B)$, $\min = \min(R, G, B)$
  - $S = (\max - \min) / \max$ (or $S = 0$, if $V = 0$)
  - $$H = 60 \times \begin{cases} 0 + (G - B) / (\max - \min), & \text{if } \max = R \\ 2 + (B - R) / (\max - \min), & \text{if } \max = G \\ 4 + (R - G) / (\max - \min), & \text{if } \max = B \end{cases}$$

  $H = H + 360$, if $H < 0$

https://blog.csdn.net/yougjiankuang/article/details/79279754

## 1.1 Image characteristic

- Colour feature
  - Feature: colour feature is a global feature
  - Methods of expressing color feature
    - Colour histogram
    - **Colour set**
    - Colour moment
    - Colour coherence vector
    - Colour correlogram



https://blog.csdn.net/qq_23291783/article/details/52036729

## 1.1 Image characteristic

- Colour feature
  - Feature: colour feature is a global feature
  - Methods of expressing color feature
    - Colour histogram
    - Colour set
    - **Colour moment**
    - Colour coherence vector
    - Colour correlogram



https://blog.csdn.net/qq_23291783/article/details/52036729

# Colour moment:

1. First-order matrix: The average intensity of color component

$$\mu_i = \frac{1}{N}\sum_{j=1}^{N} P_{ij}$$

2. Second-order matrix: The variance of color component

$$\sigma_i = \left[\frac{1}{N}\sum_{j=1}^{N} (P_{ij} - \mu_i)^2\right]^{\frac{1}{2}}$$

3. Third-order matrix: The skewness of color component

$$s_i = \left[\frac{1}{N}\sum_{j=1}^{N} (P_{ij} - \mu_i)^3\right]^{\frac{1}{3}}$$



https://blog.csdn.net/qq_23291783/article/details/52036729

## 1.1 Image characteristic

- Colour feature
  - Feature: colour feature is a global feature
  - Methods of expressing color feature
    - Colour histogram
    - Colour set
    - Colour moment
    - **Colour coherence vector**
    - **Colour correlogram**

## 1.1 Image characteristic

- Texture feature
  - **Feature: texture feature is a global feature**
  - Methods of expressing texture feature
    - Statical method
    - Geometrical method
    - Model method
    - Signal processing method

## 1.1 Image characteristic

- Texture feature

  Feature: texture feature is a global feature

  Methods of expressing texture feature

  **Statical method**
  **Geometrical method**
  Model method
  Signal processing method

## 1.1 Image characteristic



- Texture feature
  - Feature: texture feature is a global feature
  - Methods of expressing texture feature
    - Statical method
    - Geometrical method
    - **Model method**
    - **Signal processing method**

## 1.1 Image characteristic

- Shape feature

点击此处添加标题

Methods of expressing shape feature

- Contour-based method
- Region-based method

- Spatial relationship feature



## Methods of expressing spatial relationship feature:

① First, the image is automatically segmented and the object or color area is divided from the image. Then the image features are extracted and indexed according to these areas.

② The image is divided into several regular sub-blocks uniformly, and then the features of each sub-block are extracted and indexed.

Image quality indexs include resolution, color depth, image distortion and so on.

- Resolution
- Color depth
- Image distortion





https://en.Wikipedia.org/wiki/image_quality

## * Intorduction to the factors that affect the vision effect

- Contrast
- Saturation
- Sharpness

## ① Basic statistical analysis:

- Amount of information-entropy
- Gray average
- Gray variance
- Gray mode
- Gray mid-value
- Gray value region

# ① Basic statistical analysis:

- Amount of information-entropy

$$H = -\sum_{i=1}^{k} p_i log_2(p_i)$$

# ① Basic statistical analysis:

- Gray average

$$\bar{f} = \sum_{i=1}^{M} \sum_{j=1}^{N} \frac{f(i,j)}{M \cdot N}$$

# ① Basic statistical analysis:

- Gray mode
- Gray mid-value
- Gray value region

## ① Basic statistical analysis:

- Gray variance

$$S = \frac{1}{M \cdot N} \sum_{i=1}^{M} \sum_{j=1}^{N} \left[ f(i,j) - \bar{f} \right]^2$$

## ② The statistical characteristics of multidimensional image:

- covariance

$$S_{fg}^{2}=S_{gf}^{2}=\frac{1}{M\cdot N}\sum_{i=1}^{M}\sum_{j=1}^{N}\left[f(i,j)-\bar{f}\right]\left[g(i,j)-\bar{g}\right]$$

$$\Sigma = \begin{bmatrix} S_{11}^2 & S_{12}^2 & \cdots & S_{1N}^2 \\ S_{21}^2 & S_{22}^2 & \cdots & S_{2N}^2 \\ \vdots & \vdots & \ddots & \vdots \\ S_{N1}^2 & S_{N2}^2 & \cdots & S_{NN}^2 \end{bmatrix}$$

https://blog.csdn.net/xholes/article/details/53291512

② The statistical characteristics of multidimensional image:

- Correlation coefficient

$$r_{fg} = \frac{S_{fg}^2}{S_{ff} S_{gg}}$$

$$R = \begin{bmatrix} 1 & r_{12} & r_{13} & \cdots & r_{1N} \\ r_{21} & 1 & r_{23} & \cdots & r_{2N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{N1} & r_{N2} & r_{N3} & \cdots & 1 \end{bmatrix}$$

## 1.3 What is Image Quality Assessment

The meaning of image quality mainly includes image fidelity and image intelligibility. Image quality assessment aims to use computational models to measure the image quality consistently with subjective.



http://sse.tongji.edu.cn/linzhang/IQA/IQA.htm

## 1.4 Application of Image Quality Assessment

- It can be used to dynamically monitor and adjust image quality.
- It can be used to optimize algorithms and parameter settings of image processing systems.
- It can be used to benchmark image processing systems and algorithms.



http://live.ece.utexas.edu/research/Quality/intro.htm

2

# Datasets and The metrics of Image Quality Assessment algorithm

## 2.1 The introduction to the datasets

| Dataset | Reference Img. No. | Distorted Img. No. | Distortion Types | Subjects No. |
|---|---|---|---|---|
| LIVE | 29 | 779 | 5 | 161 |
| TID2008 | 25 | 1700 | 17 | 838 |
| TID2013 | 25 | 3000 | 24 | 917 |
| CSIQ | 30 | 866 | 6 | 35 |

- MOS: Mean Opinion Score
- DMOS: Differential Mean Opinion Score

## ① Root mean squared error---RMSE

- Formula:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_{observed,i} - X_{model,i})^2}{n}}$$

② Pearson product-moment correclation coefficient---PLCC

- Formula:

$$plcc = \frac{COV(X,Y)}{\delta_X \delta_Y}$$

$$= \frac{\sum(X-\bar{X})(Y-\bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i-\bar{X})^2 \sum_{i=1}^{n}(Y_i-\bar{Y})^2}}$$

③ Spearman rank order correlation coefficient---SROCC

- Formula:

$$srocc = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

④ Kendall rank order correlation coefficient---KROCC

- Formula:

$$krocc = \frac{N_c - N_d}{\frac{1}{2}N(N-1)}$$

3

# Methods

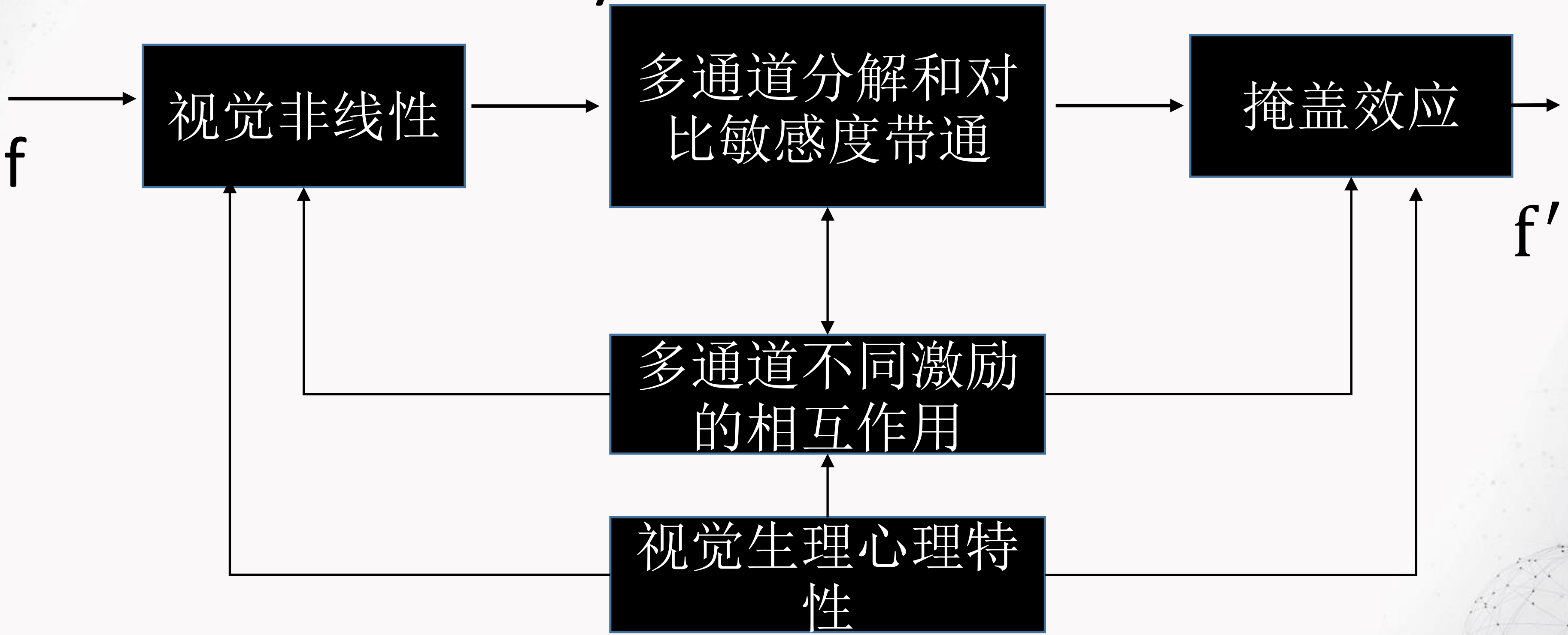
- Subjective Assessment
- Objective Assessment

- What is subjective assessment
- The advantages and disadvantages of subjective assessment
- Human Visual System

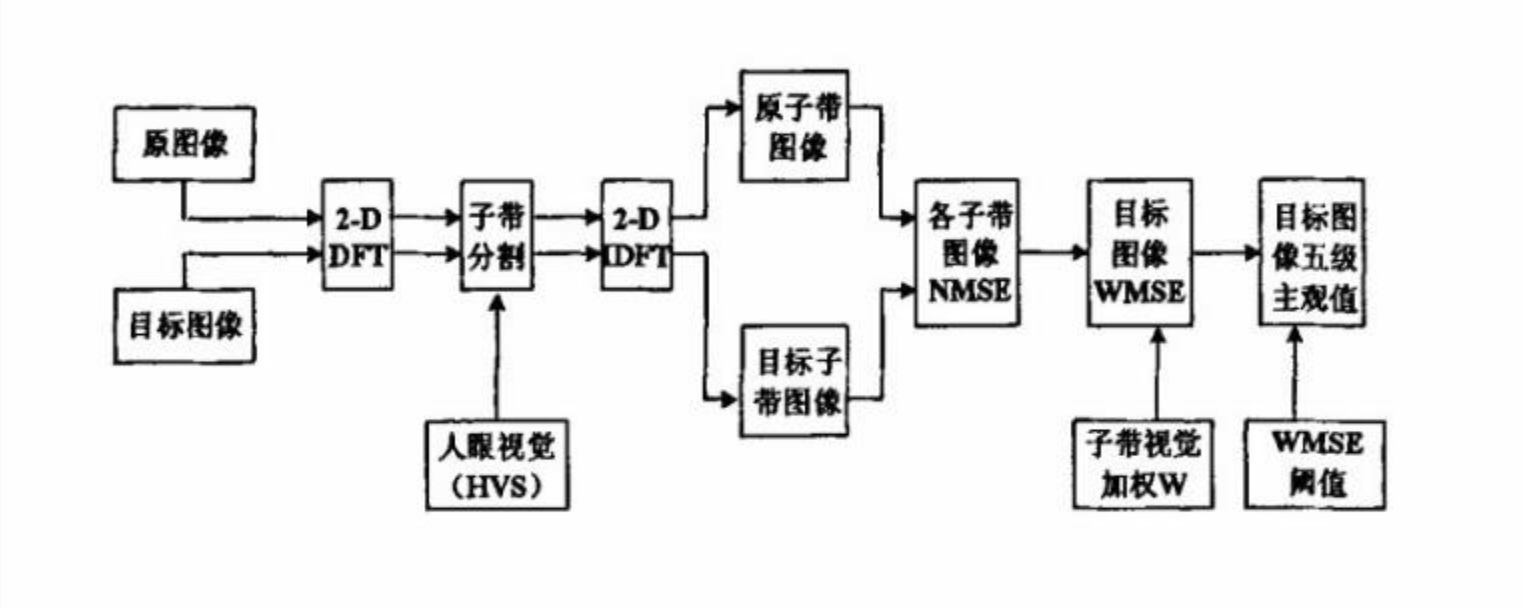
- Human Visual System

  - Visual luminance characteristic
  - Visual spatial frequency characteristic
  - Visual time frequency characteristic

- Human Visual System



# Evaluation method based on visual perception---PQS




https://max.book118.com/html/2016/1220/74878196.shtm

- What is Objective Assessment
- Full-Reference Image Quality Assessment
- Reduced-Reference Image Quality Assessment
- No-Reference Image Quality Assessment

# The basic process of the top-down approach:

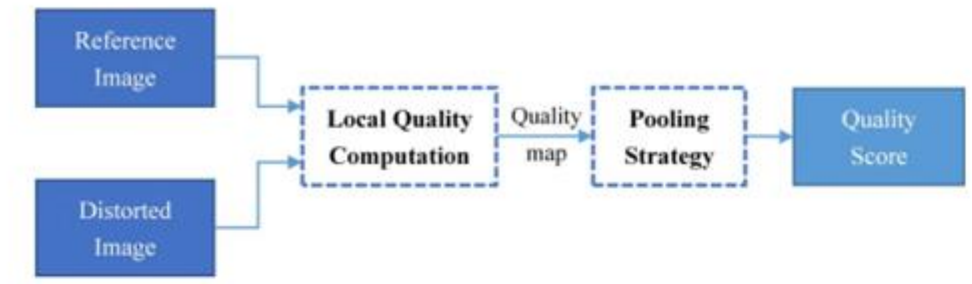



Fig. 1. The flowchart of a class of two-step FR-IQA models.



https://www.jianshu.com/p/b12dc6d634c1

# 1.PSNR(Peak Signal-to-Noise Ratio)

- Mean Square Error----MSE

$$MSE = \frac{1}{M \times N} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (f(x,y) - f'(x,y))^2$$

$$PSNR = 10 \times log \frac{L \times L}{MSE}$$



https://blog.csdn.net/majinlei121/article/details/46779125

# 3. Multi-scale structural similarity—MS-SSIM

$$SSIM(x,y) = [l_M(x,y)]^{\alpha_M} \cdot \prod_{j=1}^{M} [c_j(x,y)]^{\beta_j} \cdot [s_j(x,y)]^{\gamma_j}$$



https://ieeexplore.ieee.org/document/1292216/#full-text-section

# 4. Gradient-based Structural Similarity—GSSIM

$$s_g(x, y) = \frac{\sigma_{x'y'} + C_3}{\sigma_{x'}\sigma_{y'} + C_3}$$

$$c_g(x, y) = \frac{2\sigma_{x'}\sigma_{y'} + C_2}{\sigma_{x'}^2 + \sigma_{y'}^2 + C_2}$$

https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4107183

$$GSSIM(x, y) = [l(x, y)]^{\alpha} \cdot [c_g(x, y)]^{\beta} \cdot [s_g(x, y)]^{\gamma}$$

# Deep Learning of Human Visual Sensitivity in Image Quality Assessment Framework

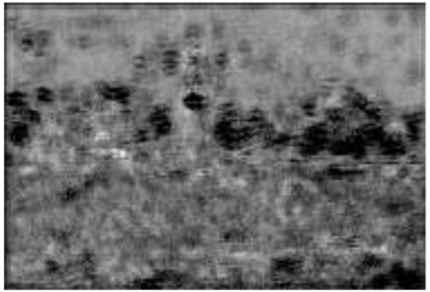
Jongyoo Kim Sanghoon Lee∗
Department of Electrical and Electronic Engineering, Yonsei Universiy, Seoul, Korea
{jongky, slee}@yonsei.ac.kr



https://ieeexplore.ieee.org/document/8099696

(a)

(b)



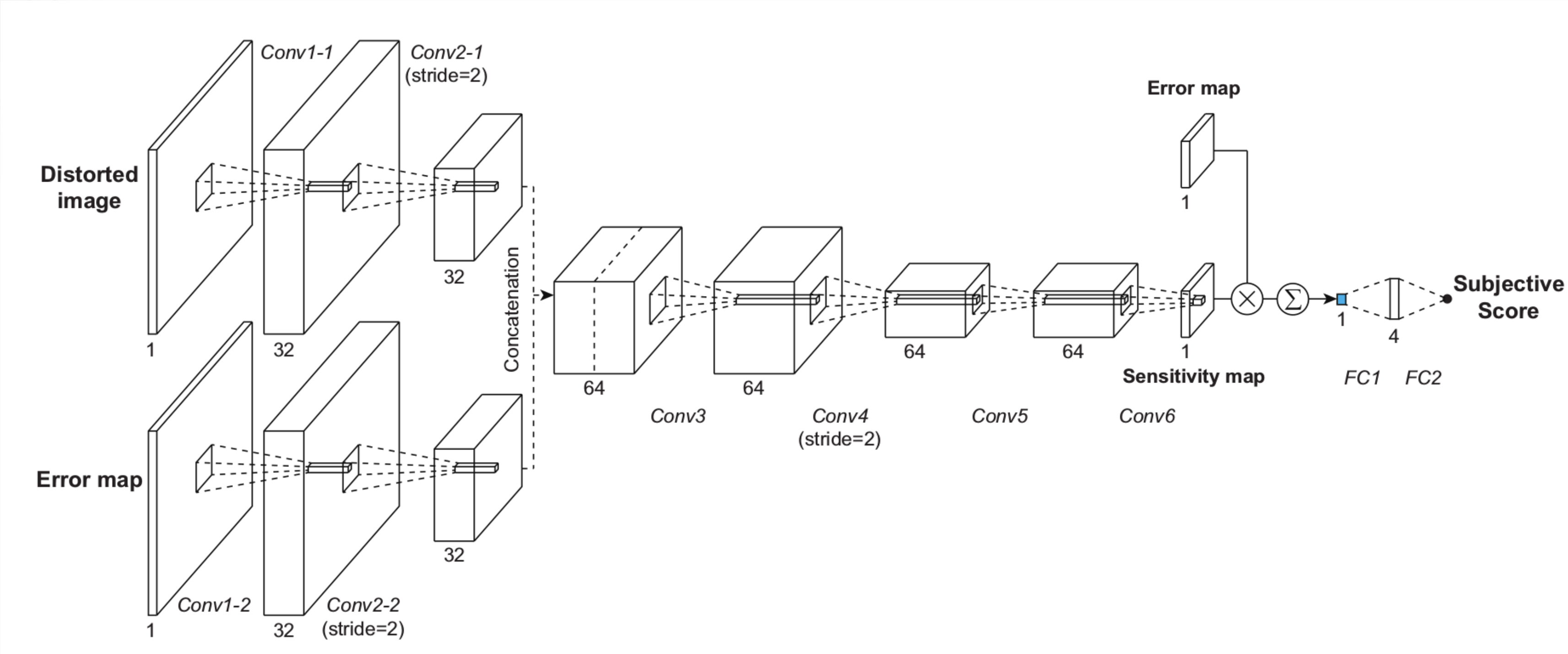
(c)

# Contributions:

- DeepQA learns the visual sensitivity characteristics of the HVS without any prior knowledge. By using a deep CNN, the visual weight of each pixel is sought by using a triplet of a distorted image, its objective error map, and its ground-truth subjective score.
- DeepQA can generate the perceptual error map as an intermediate result, which provides us an intuitive analysis of local artifacts for given distorted images.

- A novel deep CNN based FR-IQA framework is proposed. Our model can be trained by end-to-end optimization, and achieves state-of-the-art correlation with human subjective scores.

# Architecture:



# Error map:

$$\mathbf{e} = \frac{\log\left(1/\left(\left(\hat{I}_r - \hat{I}_d\right)^2 + \varepsilon/255^2\right)\right)}{\log\left(255^2/\varepsilon\right)}$$

- **Sensitivity Map Prediction:**

$$\mathbf{s}_1 = CNN_1(\hat{I}_d; \theta_1)$$

$$\mathbf{s}_2 = CNN_2(\hat{I}_d, \mathbf{e}; \theta_2)$$

- **Perceptual error map:**

$$\mathbf{P} = \mathbf{s} \odot \boldsymbol{e}$$

# Loss Function：

$$\mathcal{L}_s(\hat{I}_d;\theta) = \|(f(\mu_{\mathbf{p}}) - S)\|_F^2$$

# Total Variation Regularization:

$$TV(\mathbf{s}) = \frac{1}{H \cdot W} \sum_{(i,j)} (sobel_h(\mathbf{s})^2 + sobel_v(\mathbf{s})^2)^{\beta/2}$$

# Experimental Results:

- Dataset: LIVE, TID2008, TID2013, CSIQ

| Type | | LIVE IQA | | CSIQ | | TID2008 | | TID2013 | | Weighted Avg. | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | SRCC | PLCC | SRCC | PLCC | SRCC | PLCC | SRCC | PLCC | SRCC | PLCC |
| FR | PSNR | 0.876 | 0.872 | 0.806 | 0.800 | 0.553 | 0.573 | 0.636 | 0.706 | 0.666 | 0.704 |
| | SSIM | 0.948 | 0.945 | 0.876 | 0.861 | 0.775 | 0.773 | 0.637 | 0.691 | 0.745 | 0.767 |
| | MS-SSIM | 0.951 | 0.949 | 0.913 | 0.899 | 0.854 | 0.657 | 0.786 | 0.833 | 0.842 | 0.809 |
| | VIF | 0.963 | 0.960 | 0.920 | 0.928 | 0.749 | 0.808 | 0.677 | 0.772 | 0.765 | 0.826 |
| | GMSD | 0.960 | 0.960 | **0.957** | **0.954** | **0.891** | 0.879 | 0.804 | 0.859 | 0.867 | 0.890 |
| | FSIMc | 0.960 | 0.961 | 0.931 | 0.919 | 0.884 | 0.876 | **0.851** | **0.877** | **0.884** | **0.893** |
| | DOG-SSIMc | 0.963 | 0.966 | 0.954 | 0.943 | **0.935** | **0.937** | **0.926** | **0.934** | **0.937** | **0.940** |
| | *FR-DCNN* | **0.975** | **0.977** | - | - | - | - | - | - | - | - |
| | *DeepQA–s* | **0.977** | **0.975** | **0.957** | **0.956** | 0.878 | **0.892** | 0.766 | 0.818 | 0.848 | 0.876 |
| | *DeepQA* | **0.981** | **0.982** | **0.961** | **0.965** | **0.947** | **0.951** | **0.939** | **0.947** | **0.949** | **0.955** |
| NR | *SESANIA* | 0.934 | 0.948 | - | - | - | - | - | - | - | - |
| | *CNN* | 0.956 | 0.953 | - | - | - | - | - | - | - | - |
| | *Patchwise* | 0.960 | 0.972 | - | - | - | - | 0.835 | 0.855 | - | - |
| | *BIECON* | 0.958 | 0.960 | - | - | - | - | - | - | - | - |

# Making a "Completely Blind" Image Quality Analyzer

Anish Mittal, Rajiv Soundararajan, and Alan C. Bovik, Fellow, IEEE





https://ieeexplore.ieee.org/abstract/document/6353522

# NIQE(Making a “Completely Blind” Image Quality Analyzer)

we call the Natural Image Quality Evaluator (NIQE) is based on the construction of a “quality aware” collection of statistical features based on a simple and successful space domain natural scene statistic (NSS) model.

# NIQE Model:

Model is based on constructing a collection of “quality aware” features and fitting them to a multivariate Gaussian (MVG) model. The quality aware features are derived from a simple but highly regular natural scene statistic (NSS) model. The quality of a given test image is then expressed as the distance between a multivariate Gaussian (MVG) fit of the NSS features extracted from the test image, and a MVG model of the quality aware features extracted from the corpus of natural images.

# Multivariate Gaussian Model:

$$f_X(x_1, \ldots, x_k) = \frac{1}{(2\pi)^{k/2}|\Sigma|^{1/2}} \times \exp\left(-\frac{1}{2}(x-\nu)^T\Sigma^{-1}(x-\nu)\right)$$

$$
\begin{aligned}
D(\nu_1, \nu_2, \Sigma_1, \Sigma_2) \\
= \sqrt{\left( (\nu_1 - \nu_2)^T \left( \frac{\Sigma_1 + \Sigma_2}{2} \right)^{-1} (\nu_1 - \nu_2) \right)}
\end{aligned}
$$

# Convolutional Neural Networks for No-Reference Image Quality Assessment

Le Kang , Peng Ye, Yi Li , and David Doermann
University of Maryland, College Park, MD, USA
NICTA and ANU, Canberra, Australia
{lekang,pengye,doermann}@umiacs.umd.edu yi.li@cecs.anu.edu.au

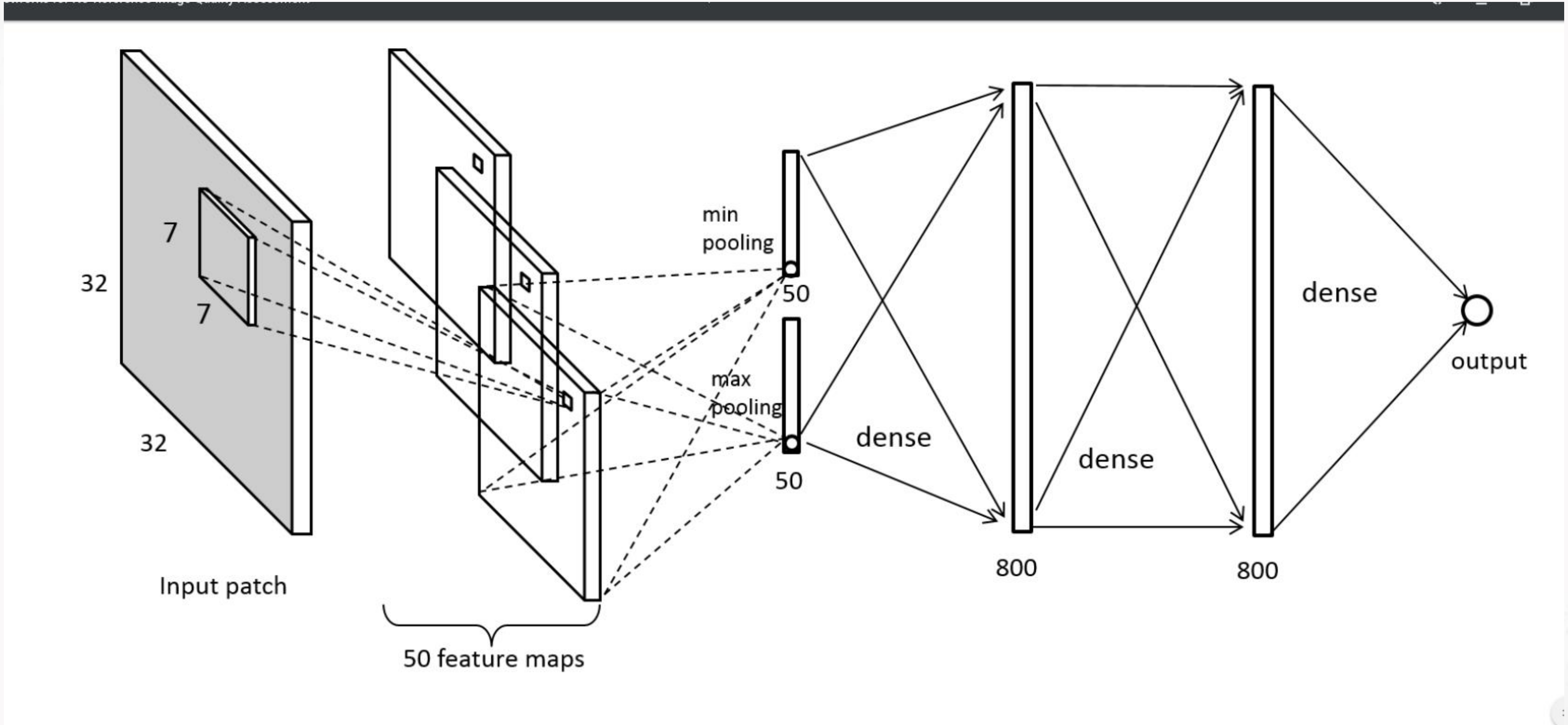


https://ieeexplore.ieee.org/document/6909620/

# Contributions:

- Author propose a novel framework that allows learning and prediction of image quality on local regions.

# Network Architecture:



# Local Normalization:

$$\hat{I}(i,j) = \frac{I(i,j) - \mu(i,j)}{\sigma(i,j) + C}$$

$$\mu(i,j) = \sum_{p=-P}^{p=P} \sum_{q=-Q}^{q=Q} I(i+p, j+q)$$

$$\sigma(i,j) = \sqrt{\sum_{p=-P}^{p=P} \sum_{q=-Q}^{q=Q} \Big(I(i+p, j+q) - \mu(i,j)\Big)^2}$$

# Pooling：

$$u_k = \max_{i,j} R_{i,j}^k$$

$$v_k = \min_{i,j} R_{i,j}^k$$

# Loss Function：

$$L = \frac{1}{N}\sum_{n=1}^{N} \| f(x_n; w) - y_n \|_{l_1}$$

$$w' = \min_{w} L$$

| SROCC | JP2K | JPEG | WN | BLUR | FF | ALL |
|---|---|---|---|---|---|---|
| *PSNR* | 0.870 | 0.885 | 0.942 | 0.763 | 0.874 | 0.866 |
| *SSIM* | 0.939 | 0.946 | 0.964 | 0.907 | 0.941 | 0.913 |
| *FSIM* | 0.970 | 0.981 | 0.967 | 0.972 | 0.949 | 0.964 |
| DIIVINE | 0.913 | 0.910 | **0.984** | 0.921 | 0.863 | 0.916 |
| BLIINDS-II | 0.929 | 0.942 | 0.969 | 0.923 | 0.889 | 0.931 |
| BRISQUE | 0.914 | 0.965 | 0.979 | 0.951 | 0.877 | 0.940 |
| CORNIA | 0.943 | 0.955 | 0.976 | **0.969** | 0.906 | 0.942 |
| CNN | **0.952** | **0.977** | 0.978 | 0.962 | **0.908** | **0.956** |

| LCC | JP2K | JPEG | WN | BLUR | FF | ALL |
|---|---|---|---|---|---|---|
| *PSNR* | 0.873 | 0.876 | 0.926 | 0.779 | 0.870 | 0.856 |
| *SSIM* | 0.921 | 0.955 | 0.982 | 0.893 | 0.939 | 0.906 |
| *FSIM* | 0.910 | 0.985 | 0.976 | 0.978 | 0.912 | 0.960 |
| DIIVINE | 0.922 | 0.921 | **0.988** | 0.923 | 0.888 | 0.917 |
| BLIINDS-II | 0.935 | 0.968 | 0.980 | 0.938 | 0.896 | 0.930 |
| BRISQUE | 0.922 | 0.973 | 0.985 | 0.951 | 0.903 | 0.942 |
| CORNIA | 0.951 | 0.965 | 0.987 | **0.968** | 0.917 | 0.935 |
| CNN | **0.953** | **0.981** | 0.984 | 0.953 | **0.933** | **0.953** |

# NIMA: Neural Image Assessent

Hossein Talebi and Peyman Milanfar

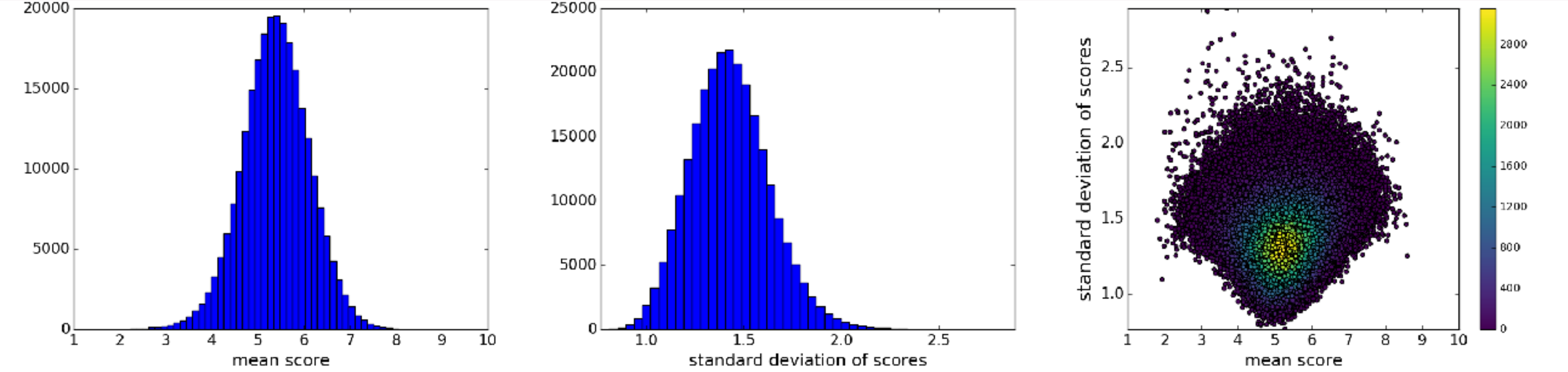
# Contributions:

- Author introduce a novel approach to predict both technical and aesthetic qualities of images.
- Author’s approach differs from others in that we predict the distribution of human opinion scores using a convolutional neural network.
- Author use the squared EMD(earth mover’s distance) loss.

# A Large-Scale Database for Aesthetic Visual Analysis(AVA)

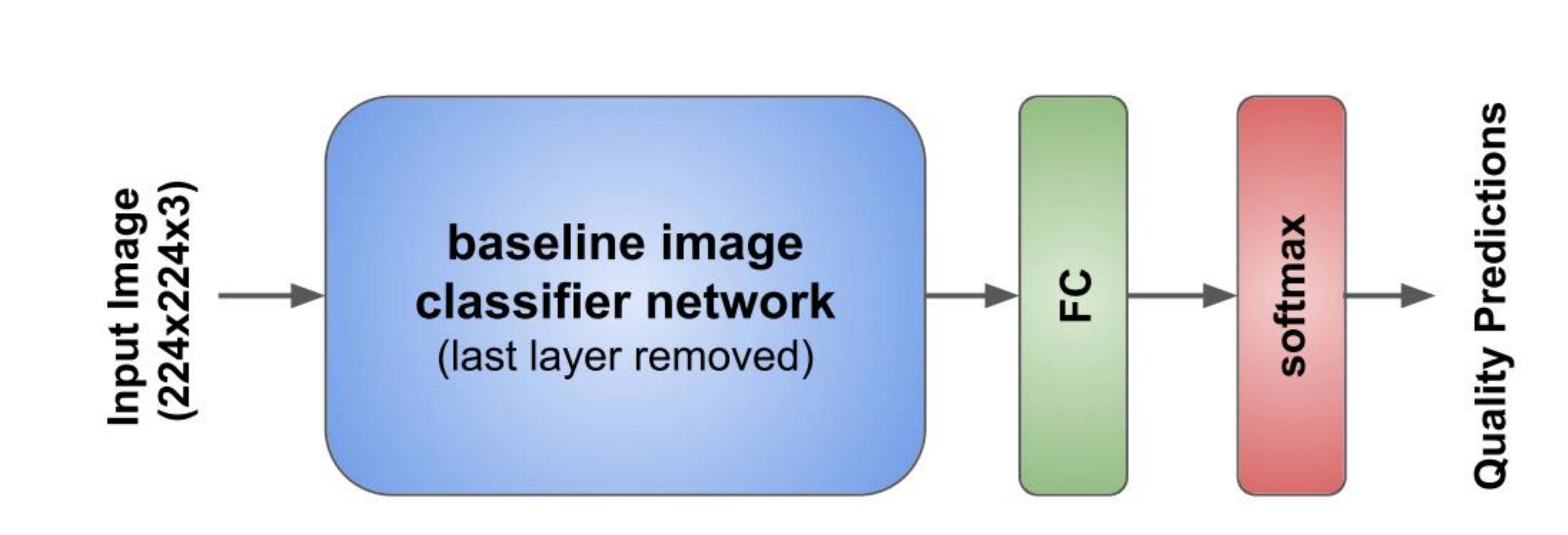
The AVA dataset contains about 255,000 images, rated based on aesthetic qualities by amateur photographers. Each photo is scored by an average of 200 people in response to photography contests. The image ratings range from 1 to 10, with 10 being the highest aesthetic score associated to an image.

# A Large-Scale Database for Aesthetic Visual Analysis(AVA)




: Histograms of ratings from AVA dataset [1]. Left: Histogram of mean scores. Middle: Histogram of standard devia

# Proposed Method



$$\mu = \sum_{i=1}^{N} s_i \times P_{s_i}$$

$$\sigma = (\sum_{i=1}^{N} (s_i - \mu)^2 \times P_{s_i})^{1/2}$$

# Loss Function:

$$\mathrm{EMD}(\mathbf{p}, \widehat{\mathbf{p}}) = \left( \frac{1}{N} \sum_{k=1}^{N} |\mathrm{CDF}_{\mathbf{p}}(k) - \mathrm{CDF}_{\widehat{\mathbf{p}}}(k)|^r \right)^{1/r}$$

(a) 6.38 (7.16)

(b) 6.24 (6.79)

(c) 6.22 (6.64)

(d) 6.16 (6.93)

(e) 5.92 (6.23)

(f) 5.71 (5.78)

(g) 5.61 (5.54)

(h) 5.28 (5.32)

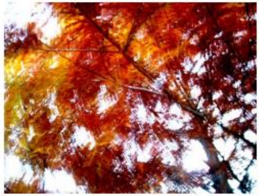
(i) 5.11 (5.23)

(j) 5.03 (5.35)

(k) 4.90 (4.91)

(l) 4.83 (4.89)

(m) 4.77 (4.55)



(n) 4.48 (3.95)

(o) 3.55 (3.53)

(a) 6.88 (7.40)

(b) 6.63 (6.89)

(c) 6.29 (6.59)

(d) 5.86 (6.16)

(e) 5.77 (5.52)


(f) 5.51 (5.47)

(g) 5.46 (5.38)

(h) 5.24 (4.74)

(i) 4.96 (4.83)

(j) 4.90 (4.71)

(k) 4.60 (4.59)

(l) 4.53 (5.05)

4

# Deep learning framework ---PyTorch

## 4.1 Brief introduction to the PyTorch

PYTORCH

Deep Learning with PyTorch

https://pytorch.org
https://pytorch.org/tutorials/beginner/deep_learning_60min_blitz.html

# Dynamic Neural Network:

A graph is created on the fly

```
from torch.autograd import Variable

x = Variable(torch.randn(1, 10))
prev_h = Variable(torch.randn(1, 20))
W_h = Variable(torch.randn(20, 20))
W_x = Variable(torch.randn(20, 10))
```

- Dynamic definition of a graph vs Static definition of a graph
- Visualization

  Tensorflow: Tensorboard

# THANKS