The Perception-Distortion Tradeoff

Yochai Blau and Tomer Michaeli

SIPL Annual Event 2018
Image Restoration

Super-resolution

Inpainting

Dehazing

Denoising

Deblurring
Image Restoration

Goals

✓ Similar to ground truth image (low distortion)
✓ Good perceptual quality

Our work

Algorithms cannot achieve both low distortion and good perceptual quality
Empirical Evidence

Ma et al. 2017

Better quality

Less distortion

No-reference quality measure

RMSE
Empirical Evidence

No-reference quality measure

Better quality

Less distortion

RMSE

SRCNN, A+
LapSRN, SelfEx
EDSR, VDSR, Bae
SRResNet, MSE
SRResNet, VGG2,2
Deng, SRGAN, MSE
SRGAN, VGG2,2, Johnson
SRGAN, VGG5,4, ENet
Empirical Evidence
Empirical Evidence

No-reference quality measure

RMSE
Empirical Evidence

No-reference quality measure

RMSE

SSIM

MS-SSIM

IFC

VIF

VGG\_2,2

SRCNN
LapSRN
EDSR
VDSR
SRResNet
MSE

A+
SelfEx
Deng
Johnson
SRGAN
VGG\_2,2
SRGAN
VGG\_5,4
Enet
Mehrrez

2VGG

2.7

2.1

3.1

2.5

2.8

3.4

0.31

0.28

0.25

0.34

0.75

0.71

0.67

0.63

0.96

0.95

0.94

0.93

2.4

2.1
Our work

- Alg. 1
- Alg. 2
- Alg. 3

Possible

Impossible

Less distortion

Distortion

Perception

Better quality
Image Restoration

Natural image \( x \sim p_X \) → Degraded \( y \) → Algorithm \( p_{\hat{X}|Y} \) → Reconstructed \( \hat{x} \)
Distortion

\[ x \sim p(x) \]

Natural image

\[ \mathbb{E}[\Delta(X, \hat{X})] \]

Reconstructed

SSD, SSIM, MS-SSIM, IFC, VIF, VGG, …
Perceptual Quality

\[ x \sim p_x \]

Natural image

50%

Real or Fake?

\[ \hat{x} \sim p_{\hat{x}} \]

Reconstructed

\[ p_{\text{success}} \propto d_{\text{TV}}(p_x, p_{\hat{x}}) \]
Perceptual Quality

\[ x \sim p_x \]

Natural image

\[ \hat{x} \sim p_{\hat{x}} \]

Reconstructed

Perceptual quality

\[ d(p_x, p_{\hat{x}}) \]
Problem Setting

Original  Degraded  Reconstructed

$X \rightarrow Y \rightarrow \hat{X}$

Distortion: $\mathbb{E}[\Delta(X, \hat{X})]$  
Perception: $d(p_X, p_{\hat{X}})$
Toy Example

\[ Y = X + N \sim N(0, 1) \]

MSE distortion:
\[ \Delta(x, \hat{x}) = (x - \hat{x})^2 \]
Toy Example

\[ Y = X + N \]

0 – 1 distortion:

\[ \Delta(x, \hat{x}) = \begin{cases} 
0 & \hat{x} = x \\
1 & \hat{x} \neq x 
\end{cases} \]
MNIST Example

Noisy ($\sigma = 1$)$

MMSE

MAP

$\sigma = 1$$$

$\sigma = 3$$$

$\sigma = 5$
The Perception-Distortion Function

\[ P(D) = \min_{p_{\hat{X}|Y}} d(p_X, p_{\hat{X}}) \quad \text{s.t.} \quad \mathbb{E}[\Delta(X, \hat{X})] \leq D \]

[Diagram showing the relationship between perception, distortion, and the minimization of distance.]
The Perception-Distortion Function

\[ P(D) = \min_{p_{\hat{X} \mid Y}} d(p_X, p_{\hat{X}}) \text{ s.t. } \mathbb{E}[\Delta(X, \hat{X})] \leq D \]

**Theorem 1**
For all popular \( d(p, q) \) and any \( \Delta(x, \hat{x}) \)
\( P(D) \) is:
1. monotonically non-increasing
2. convex
The Perception-Distortion Function

\[ P(D) = \min_{p_{\hat{X}|Y}} d(p_X, p_{\hat{X}}) \text{ s.t. } \mathbb{E}[\Delta(X, \hat{X})] \leq D \]

**Theorem 2**
For the RMSE distortion, 
\[ \frac{D_{\text{max}}}{D_{\text{min}}} \leq \sqrt{2} \]
Empirical Evidence

No-reference quality measure

RMSE

-3 dB
Connection to Rate-Distortion

\[ P(D) = \min_{p_{\hat{X}|Y}} d(p_X, p_{\hat{X}}) \quad \text{s.t.} \quad \mathbb{E}[\Delta(X, \hat{X})] \leq D \]

\[ R(D) = \min_{p_{\hat{X}|X}} I(X, \hat{X}) \quad \text{s.t.} \quad \mathbb{E}[\Delta(X, \hat{X})] \leq D \]
Traversing the Tradeoff

\[ P(D) = \min d(p_X, p_{\hat{X}}) \quad \text{s.t.} \quad \mathbb{E}[\Delta(X, \hat{X})] \leq D \]

\[ \ell_{\text{gen}} = \ell_{\text{distortion}} + \lambda \ell_{\text{adv}} \approx \mathbb{E}[\Delta(X, \hat{X})] + \lambda d(p_X, p_{\hat{X}}) \]

GAN-Based Image Restoration
- Ledig et al., CVPR ’17
- Yeh et al., CVPR ’17
- Sajjadi et al., ICCV ’17
- ...
Traversing the Tradeoff with a GAN
I. A tradeoff between perceptual quality and distortion

II. The “optimal algorithm” is application dependent

III. After choosing a distortion level, move towards the bound

IV. GANs allow to systematically approach the bound