High Dynamic Range Imaging using Deep Image Priors

ICASSP 2020 - Paper ID 4320

Gauri Jagatap and Chinmay Hegde

New York University
Tandon School of Engineering

Motivation: HDR Imaging

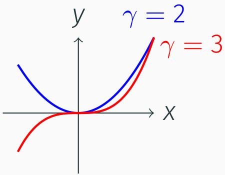
Low light imaging

- Limited camera sensor hardware and high photon noise can result in images with low dynamic range resolution.
- Goal: Novel techniques for improved high dynamic range (HDR) images from camera sensor data.

Models for HDR: Gamma encoding

- Low-light image acquisition can be viewed as a non-linear forward problem where each "true pixel intensity" is distorted.
- Low-light images are also corrupted by (additive) photon sensor noise, so that the effects of this noise are amplified in a non-linear manner when gamma correction is applied.
- Forward model:

$$f(x) = x^{\gamma} + \epsilon$$



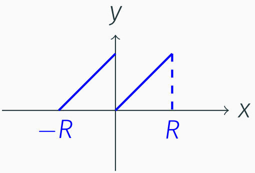
$$f(x) = x^{\gamma}$$

Figure 1: Gamma encoding

Models for HDR: Modulo sensing

- A modulo camera sensor folds the pixel intensities into an interval via a sawtooth transfer function.

- Whenever the pixel value of the camera sensor saturates, the pixel counter is reset to zero and photon collection continues till the next saturation point.



$$f(x) = \mod(x, R)$$

Figure 2: Modulo sensing

- Task of inverting modulo-sensed images is highly ill-posed.

Inverse imaging

Given a d-dimensional image signal x^* and a sensing operator $f(\cdot) : \mathbb{R}^d \to \mathbb{R}^n$, measurements y take the form:

$$y = f(x^*)$$

Task: Recover x^* from measurements y.

- Posed as an optimization problem:

$$\hat{x} = \arg\min_x L(x) = \arg\min_x \|y - f(x)\|_2^2$$

- d-dimensional input image requires $n = O(d)$ measurements in conventional sensing systems for stable estimation.
- f can in general be ill-posed.

Structured image recovery

Denoising

Inpainting

Super-res

- Degrees of freedom of natural images is typically lower than d.
- Constrain the search space to this lower-dimensional set $\mathcal{S}$.

$$\hat{x} = \arg\min_{x \in \mathcal{S}} L(x) = \arg\min_{x \in \mathcal{S}} \|y - f(x)\|_2^2$$

- Examples of $\mathcal{S}$: sparsity, total variation, dictionary models, neural generative models, ...

Structure: Deep image prior (DIP)

Our contributions[1]:

- Deep image prior for inverting HDR imaging models:
    - Gamma encoding.
    - (Compressive) modulo sensing.

[1] G. Jagatap and C. Hegde, ”High Dynamic Range Imaging using Deep Image Priors,” ICASSP (2020).

Structure: Deep image prior (DIP)

Our contributions[1]:

- Deep image prior for inverting HDR imaging models:
  - Gamma encoding.
  - (Compressive) modulo sensing.

Prior $\mathcal{S}$	Training data?	Neural ?
Sparsity (w or w/o structure, total variation)	No	No
Autoencoders	Yes	Yes
Deep learned generative priors	Yes	Yes
Deep image prior	No	Yes

Table 1: Low-dimensional priors

[1] G. Jagatap and C. Hegde, "High Dynamic Range Imaging using Deep Image Priors," ICASSP (2020).

Reconstruction algorithms

Formulation

Consider the HDR image recovery problem where measurements $y = f(x^*)$, and the forward transfer function f is one of the below two forms:

- Noisy gamma encoding : $y = x^{*\gamma} + \epsilon$
- Compressive modulo sensing (restricted to two periods): $y = \textbf{mod}\ (Ax^*, R)$

where $\epsilon \sim \mathcal{N}(0, \sigma^2)$, $\mathbf{1}_{Ax^*<R}$ is an element-wise indicator, $x^* \in \mathbb{R}^d$, $y \in \mathbb{R}^n$, and entries of A are from $\mathcal{N}(0, 1/n)$ with $n < d$.

Formulation

Consider the HDR image recovery problem where measurements $y = f(x^*)$, and the forward transfer function f is one of the below two forms:

- Noisy gamma encoding : $y = x^{*\gamma} + \epsilon$
- Compressive modulo sensing (restricted to two periods): $y = \bmod\ (Ax^*, R)$

where $\epsilon \sim \mathcal{N}(0, \sigma^2)$, $\mathbf{1}_{Ax^* < R}$ is an element-wise indicator, $x^* \in \mathbb{R}^d$, $y \in \mathbb{R}^n$, and entries of A are from $\mathcal{N}(0, 1/n)$ with $n < d$.

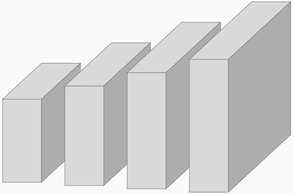
Pose recovery as the following optimization problem:

$$\min_{x,\mathbf{w}} \|y - f(x)\|_2^2 \quad \text{s.t.} \quad x = G(\mathbf{w}, z) \in \mathcal{S}$$

where S captures *Deep Image Prior* (DIP).

DCGAN Prior

A given image $x \in \mathbb{R}^{d \times k}$ is said to obey a deep decoder prior if it belongs to a set $\mathcal{S}$ defined as: $\mathcal{S} := \{x | x = G(\mathbf{w}; z)\}$ where z is a (randomly chosen, fixed) latent code vector and $G(\mathbf{w}; z)$ has the form as below.



$$x = G(\mathbf{w}, z) = \sigma(W_L \circledast \sigma(W_{L-1} \ldots \sigma(W_1 \circledast Z_1))))$$

$\sigma(\cdot)$ represents ReLU, W_i's are transposed convolutional filters and $\circledast$ represents transposed convolutional operation.

Gamma Correction and Denoising with DIP

$$\min_{x \in \mathcal{S}} L(x) = \min_{x \in \mathcal{S}} \|y - x^\gamma\|_2^2$$

Gamma Correction and Denoising with DIP

$$\min_{x \in \mathcal{S}} L(x) = \min_{x \in \mathcal{S}} \|y - x^{\gamma}\|_2^2$$

$$\min_{w} \mathcal{L}(w) = \min_{w} \|y - G(w; z)^{\gamma}\|_2^2 + \lambda \mathrm{TV}(G(w; z))$$

Algorithm 2 Gamma decoding with DIP.

1: **Input:** $z = \mathrm{vec}(Z_1), \eta, \mathrm{w}^0$.
2: **while** termination condition not met **do**
3: $\quad \mathrm{w}^{t+1} \leftarrow \mathrm{w}^{t} - \eta \nabla \mathcal{L}(\mathrm{w}^{t})$ {gradient step}
4: **end while**
5: **Output** $\hat{x} \leftarrow G(\mathrm{w}^{T}; z)$.

Modulo reconstruction with DIP

Solve: $\min_{x \in \mathcal{S}} L(x) = \min_{x \in \mathcal{S}} \|y - \mod(Ax, R)\|_2^2$

Algorithm 3 Modulo sensing with Deep Image Prior.

INITIALIZATION STAGE

Input: $A, z = \text{vec}(Z_1), \eta, \mathbf{w}^0$.

$y_{init} = y - R \cdot p_{init}$

$x^0 = \arg\min_{x \in \mathcal{S}} \|y_{init} - Ax\|_2^2$ {Net-GD for CS}

DESCENT STAGE

Input: $A, z, x^0, \eta, \mathbf{w}^0$

while termination condition not met **do**

  $p^t = \mathbf{1}_{Ax^t < 0}$

  $y_c = y - R \cdot p^t$

  $v^t \leftarrow x^t - \eta \nabla_x \|y_c - Ax\|_2^2$ {gradient step}

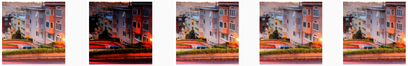
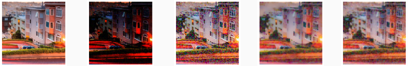
  $x^{t+1} \leftarrow \arg\min_{x \in \mathcal{S}} \|v^t - x\|_2^2$ {project to $\mathcal{S}$}

end while

Output $\hat{x} \leftarrow G(\mathbf{w}^T; z)$.

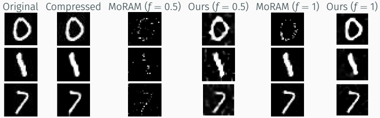
Experiments

Gamma correction and denoising

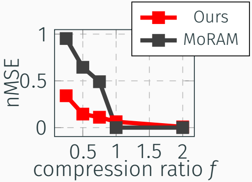
γ, σ	Original	Dark+Noise	GC	GC+TV	GC+DIP
$3, 0.01$					
	SNR (dB)	10.03	26.41	27.87	**28.33**
	SSIM	0.4657	0.9243	0.9338	**0.9413**
$4, 0.03$					
	SNR (dB)	8.56	16.77	21.44	**21.55**
	SSIM	0.2952	0.6364	0.7445	**0.7748**
$3, 0.01$					
	SNR (dB)	10.16	25.33	27.89	**28.14**
	SSIM	0.3994	0.8656	0.9038	**0.9108**
$4, 0.03$					
	SNR (dB)	8.78	16.18	**22.68**	22.64
	SSIM	0.2371	0.4914	0.7482	**0.7518**
	(a)	(b)	(c)	(d)	(e)

(a) Original image, (b) image darkened with factor γ, followed by addition of noise with variance σ, (c) gamma corrected image (d) gamma correction followed by TV denoising (e) Algorithm 1 using Deep Image Prior.

Reconstruction from compressive modulo measurements



(a)



(b)

Reconstructed images from modulo measurements (a) at compression rates of $f = n/d = 0.5, 1$ for MNIST images, (b) nMSE at different compression rates $f = n/d$ for MNIST digit '1' averaged over 10 trials, and comparison with sparsity based modulo inversion (MoRAM)

Conclusions and future directions

- Our contributions:
  - Novel applications of untrained neural priors to HDR imaging:
    (i) gamma correction and denoising;
    (ii) reconstruction from modulo observations.
  - superior empirical performance.
- Future directions:
  - Theoretical guarantees for HDR inverse problems.
  - Use of invertible neural architectures for HDR reconstruction.

Paper:
https://gaurijagatap.github.io/assets/hdrimage.pdf