Posterior Sampling Using Diffusion Models for HDR Reconstruction

Stanford | ENGINEERING
Electrical Engineering

Jamin Xie

Computational Imaging Lab, Department of Electrical Engineering, Stanford University

Introduction

- **Inverse problems:** recover unknown signal $\boldsymbol{x}_0 \in \mathbb{R}^n$ from noisy measurements $\boldsymbol{y} = \mathcal{A}(\boldsymbol{x}_0) + \boldsymbol{n} \in \mathbb{R}^m$, where $m < n$
- **Bayesian framework:** sample from posterior distribution $p(\boldsymbol{x}_0|\boldsymbol{y}) \propto p(\boldsymbol{y}|\boldsymbol{x}_0)p(\boldsymbol{x}_0)$ to get solution
- **Nonlinear HDR imaging problem:** convert Low Dynamic Range (LDR) images to High Dynamic Range (HDR) to capture wider intensity levels and improve visual realism
- **Posterior sampling & diffusion models:** leverage powerful pre-trained diffusion models (DMs) as priors combined with measurements to do posterior sampling for HDR reconstruction
- **Inference only**: no model training or fine-tuning required
- **Latent space:** working in latent space (like LDMs) using encoder and decoder could improve flexibility, efficiency, scalability, and inference time compared to pixel-based methods

Problem Statement

- **Single-exposure:** expanding dynamic range from single noisy and degraded measurement avoids alignment issues and motion artifacts that plague multi-exposure methods
- **Goal:** generate LDR with higher dynamic range using input simulated measurements distorted by known forward operator, task becomes increasing dynamic range by 2x
- **LDR datasets:** evaluate on FFHQ 256 × 256 and ImageNet 256 × 256 using 10 images from each test set

Sample trajectory of DAPS on FFHQ
First row: $\boldsymbol{x}_{0|\boldsymbol{y}}$ Second row: $\hat{\boldsymbol{x}}_0(\boldsymbol{x}_t)$ Third row $\boldsymbol{x}_t$
Left to right, the noise anneals from σ_T to 0. Row 3 to 2 is reverse diffusion. Row 2 to 1 is Langevin dynamics. Row 1 to 3 is forward diffusion.

Methods

Diffusion models

Continuous-time SDE: $t = 0$ is data, $t = T$ is Gaussian noise

Forward SDE: $\mathrm{d}\boldsymbol{x}_t = \boldsymbol{f}(\boldsymbol{x}_t, t)\mathrm{d}t + g(t)\mathrm{d}\boldsymbol{W}_t$

Reverse SDE: $\mathrm{d}\boldsymbol{x}_t = \left(\boldsymbol{f}(\boldsymbol{x}_t, t) - g^2(t)\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t)\right)\mathrm{d}t + g(t)\mathrm{d}\boldsymbol{W}_t$

Variance exploding (VE) SDE: $\boldsymbol{f}(\boldsymbol{x}_t, t) = \boldsymbol{0}$, $g(t) = \sqrt{2\dot{\sigma}_t\sigma_t}$

Sample from data dist. $p(\boldsymbol{x}_0)$: draw from $\mathcal{N}(\boldsymbol{0}, \sigma_T^2\boldsymbol{I})$, then use score $\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t)$ from pre-trained diffusion model and reverse SDE

Bayesian framework for posterior sampling

Modified VE reverse SDE using Bayes' Theorem:

$$\mathrm{d}\boldsymbol{x}_t = -2\dot{\sigma}_t\sigma_t\nabla_{\boldsymbol{x}_t}\log p_t(\boldsymbol{x}_t)\mathrm{d}t - 2\dot{\sigma}_t\sigma_t\nabla_{\boldsymbol{x}_t}\log p_t(\boldsymbol{y}|\boldsymbol{x}_t)\mathrm{d}t + \sqrt{2\dot{\sigma}_t\sigma_t}\mathrm{d}\boldsymbol{W}_t$$

Decoupled Annealing Posterior Sampling (DAPS)

Uses noise annealing process, iteratively sampling from $p(\boldsymbol{x}_t|\boldsymbol{y})$ with noise decreasing from σ_T to 0. Consecutive samples are decoupled (conditionally independent). Repeat steps 1-3 to get $\boldsymbol{x}_0$:

1. Reverse diffusion: estimate posterior mean $\hat{\boldsymbol{x}}_0(\boldsymbol{x}_t)$ using Euler ODE solver, starting from $\boldsymbol{x}_t$, plugging in score from pre-trained DM: $p(\boldsymbol{x}_0|\boldsymbol{x}_t) \approx \mathcal{N}(\boldsymbol{x}_0; \hat{\boldsymbol{x}}_0(\boldsymbol{x}_t), r_t^2\boldsymbol{I})$
2. Langevin dynamics: sample $\boldsymbol{x}_{0|\boldsymbol{y}}$ using $\hat{\boldsymbol{x}}_0(\boldsymbol{x}_t)$, forward operator: $\boldsymbol{x}_{0|\boldsymbol{y}} \sim p(\boldsymbol{x}_0|\boldsymbol{x}_t, \boldsymbol{y}) \propto p(\boldsymbol{x}_0|\boldsymbol{x}_t)p(\boldsymbol{y}|\boldsymbol{x}_0)$
3. Forward diffusion: sample $\boldsymbol{x}_t \sim \mathcal{N}(\boldsymbol{x}_{0|\boldsymbol{y}}, \sigma_t^2\boldsymbol{I})$

Conclusion and Future Work

- DAPS corrects global errors by decoupling consecutive sample points in diffusion sampling trajectory (naive reverse SDE solver takes small steps resulting in only local error corrections)
- Pixel-space DAPS performed better and faster than LatentDAPS
- Posterior sampling using pre-trained diffusion models leverages powerful priors and does not require additional training, applicable to nonlinear inverse problems like HDR reconstruction
- **Future work:** true HDR using paired LDR-HDR data, using Poisson noise for underexposed scenes

Results