NoiseCtrl: A Sampling-Algorithm-Agnostic Conditional Generation Method for Diffusion Models

Longquan Dai, He Wang, Jinhui Tang
Nanjing University of Science and Technology

Introduction:

Diffusion Model Denoise Step:

$$\mathbf{x_{t-1}} = \boldsymbol{\alpha_t}\mathbf{x_t} - \boldsymbol{\beta_t}\boldsymbol{\varepsilon_\theta}(\mathbf{x_t}, \mathbf{t}, \mathbf{c}) + \boldsymbol{\sigma_t}\boldsymbol{\varepsilon}$$
$$\mathbf{x_{t-1}} = \boldsymbol{\mu_t} + \boldsymbol{\sigma_t}\boldsymbol{\varepsilon}$$

- **TL;DR:**
  Control Diffusion Models via **Random Item**

- **Task:**
  Conditional Image Generation for Diffusion Models, which aiming to add spatially localized input conditions (eg. Depth Map, Canny Edge) to a pretrained text-to-image diffusion model.

- **Existing Work:**
  Existing methods usually control diffusion models via modifying **predicted item**.
  - Training-Required Methods: require additional training data and computational resources
  - Training-Free Methods: disrupt the stability of the sampling scheduler and requires extensive experimentation to determine optimal settings.

- **Our Method:**
  NoiseCtrl combine **predicted mean value μ_t** with **conditional noise ε**, producing a result that is attentive to the given condition.

Method:

NoiseCtrl

$$\mathbf{x_{t-1}} = \boldsymbol{\mu_t} + \boldsymbol{\sigma_t}\,\boldsymbol{\varepsilon'_t}$$

Conditional Noise

- **Conditional Noise:**

$$\boldsymbol{\varepsilon'_t} = \mathrm{r\,u}$$

where r is radius and u is direction, r and u are sampled from the joint distribution $\int_{r,u;d,\kappa}(r,u;d,\kappa)$

$$\int_{r,u;d,\kappa}(r,u;d,\kappa) = f_r(r) \cdot f_{u;d,\kappa}(u;d,\kappa)$$

- **Radius r:**

$$\mathrm{r} \sim \mathrm{p_r(r)} = \frac{\mathrm{r^{n-1}exp(-\frac{r^2}{2})}}{2^{\frac{n}{2}-1}\Gamma(\frac{n}{2})}$$

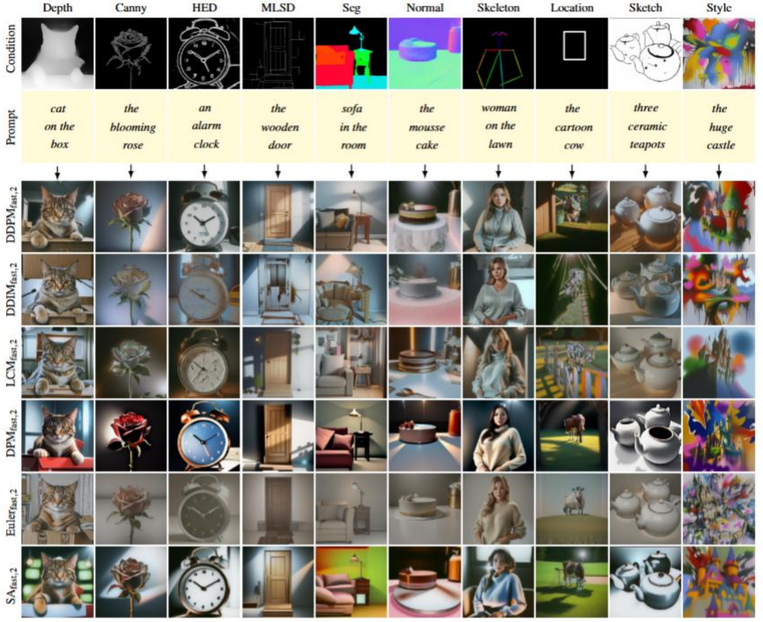
- **Direction u:**

$$\boldsymbol{d} = -\frac{\nabla_{z'}\,\mathcal{L}'_y(z')}{\left\|\nabla_{z'}\,\mathcal{L}'_y(z')\right\|_2}$$

- By simply replacing the noise component with the noise generated by our method, NoiseCtrl can be adapted to:

DDPM | DDIM | LCM | DPM | Euler | SA-Solver

Experiment:



Conclusion

- We introduce the use of conditional noise in place of Gaussian noise for coherent generation without altering existing sampling algorithms.

- We present both fast and slow direction vector estimation methods to determine the condition of the conditional noise, and we advocate for the use of the von Mises-Fisher distribution instead of the uniform distribution for expedited sampling.